NATURESERVE REPORT • DECEMBER 2011

ASSESSMENT OF WETLAND ECOSYSTEM CONDITION ACROSS LANDSCAPE REGIONS

A MULTI-METRIC APPROACH

A Report to the Environmental Protection Agency

NatureServe

NatureServe is a non-profit organization dedicated to providing the scientific basis for effective conservation.

Citation: Faber-Langendoen, D., C. Hedge, M. Kost, S. Thomas, L. Smart, R. Smyth, J. Drake, and S. Menard, 2011. *Assessment of wetland ecosystem condition across landscape regions: a multi-metric approach.* NatureServe, Arlington, VA. + Appendices.

with contributions to Section A from:

Joe Rocchio, Gwen Kittel, Pat Comer, Lesley Sneddon, and Milo Pyne

Cover Photo: Southern Wet Meadow, Bald Mountain Recreation Area, Oakland Co., Mich. Steve Thomas (EO 8240, Site 96).

Funding for this report was provided by a grant from the Environmental Protection Agency (*www.epa.gov*) through Cooperative Agreement No. RM-83377501. Although EPA made comments on and suggestions on the document, the views expressed are those of NatureServe. EPA does not endorse any products or commercial services mentioned in this publication.

NatureServe
4600 N. Fairfax Drive, 7th Floor.
Arlington, VA 22203
703-908-1800
www.natureserve.org

ASSESSMENT OF WETLAND ECOSYSTEM CONDITION ACROSS LANDSCAPE REGIONS

A MULTI-METRIC APPROACH

Don Faber-Langendoen, Cloyce Hedge, Mike Kost, Steve Thomas,
Lindsey Smart, Regan Smyth, Jim Drake, and Shannon Menard

NatureServe
A Network Connecting Science With Conservation

December 2011

ACKNOWLEDGMENTS

Developing the overall methodology for ecological assessments has been a team effort over many years. The groundwork for this report was provided by the co-authors of Faber-Langendoen et al. (2008). Their work, in turn, drew from NatureServe's Ecological Integrity Assessment Working Group. Members consisted of NatureServe and network member program staff, including Don Faber-Langendoen and Pat Comer (co-chairs), and David Braun, Elizabeth Byers, John Christy, Greg Kudray, Gwen Kittel, Shannon Menard, Esteban Muldavin, Milo Pyne, Carl Nordman, Joe Rocchio, Mike Schafale, Lesley Sneddon, and Linda Vance. More recently, the overall framework for the methodology has been summarized by Unnasch et al. (2009), and we have drawn from some of that publication for this report.

This final report was developed under Cooperative Agreement No. RM-83377501 awarded by the U.S. Environmental Protection Agency (EPA). Although EPA made comments and suggestions on the document, the views expressed are those of NatureServe. EPA does not endorse any products or commercial services mentioned in this publication. We are grateful for support from staff of the Environmental Protection Agency, including the project management provided by Anett Trebitz, and ongoing support from Rich Sumner and Mike Scozzafava for their support, and for the peer review comments from Rich Sumner, Peter Jackson and Sue Elston.

The project was undertaken by staff at NatureServe, with field crews coordinated by the Michigan Natural Features Inventory (MNFI) and the Indiana Natural Heritage Program (INNHP). We thank our field staff. In Indiana, they included Roger Hedge, Mike Homoya, Tom Post, Tom Swinford, Rich Dunbar, Nate Simons, Derek Nimetz, and Jason Larson. Office support, including compiling site information and data entry, was provided by Ron Hellmich and Breanna Sowers. In Michigan, field staff included John Fody, Aimee Kay, Jeff Lee, Monique Gorecki, Jesse Lincoln, and Aaron Kortenhoven. Office staff support was provided by Rebecca Rogers, Helen Enander, Suzanne Ridge, and Nancy Toben.

Critical database support was provided by Kristin Snow and Mary Harkness. They helped design and implement the Ecological Observations database, which houses all information collected during the project, and which can be linked to the Biotics databases used by the network member programs. We thank them for their efforts.

References

Faber-Langendoen, D., G. Kudray, C. Nordman, L. Sneddon, L. Vance, E. Byers, J. Rocchio, S. Gawler, G. Kittel, S. Menard, P. Comer, E. Muldavin, M. Schafale, T. Foti, C. Josse,and J. Christy. 2008. *Ecological Performance Standards for Wetland Mitigation based on Ecological Integrity Assessments*. NatureServe, Arlington, VA. + Appendices.

Table of Contents

Executive Summary

Many ecosystem monitoring and assessment programs are expanding their focus to address changes in ecosystem condition. This is a challenging task, given the complexity of ecosystems and the changes they undergo in response to a variety of human activities and landscape alterations. Agencies and organizations are in need of ecological methods that can address this aspect of ecosystems. These methods must be sensitive to variations in ecosystem type, size, and landscape setting, along with the dramatic losses and degradation that have occurred.

NatureServe, in partnership with member programs from the natural heritage network and federal agencies, has developed an assessment of ecosystems condition, structured around the concept of ecological integrity. Here we first (section A) review and update the overall conceptual model, and second (Section B), develop a sampling design for identifying a suite of 277 sites, primarily from network program databases. These sites span all wetland type and conditions in southern Michigan and northern Indiana. Third (Section C), we test the method on the 277 sites to determine whether it could accurately distinguish ranges of integrity across all wetland types. We tested and applied the method using a multi-level framework (remote, rapid, intensive) and multiple metrics that cover hydrology, soils, vegetation, size, buffer, and landscape. Data were summarized using an overall Index of Ecological Integrity and scorecard, focusing on rapid field assessments scores.

Main Objectives

A. **Develop a methodology for assessing wetland condition based on ecological integrity**. A conceptual model was developed to facilitate identification of a scientifically defensible set of metrics, at multiple scales. The methodology informs two main areas of wetland inventory and monitoring:

- *Baseline inventory and ambient monitoring of wetland condition*. Our methodology is applicable for local, state, and national inventory and monitoring of wetlands over time at a variety of levels (remote, rapid, intensive).
- *Mitigation and restoration of wetlands.* Our methodology sets ecological performance standards to assess site-specific and watershed-scale mitigation and restoration projects. Our methods complement other rapid assessments that are strictly mitigation focused, such as the Michigan Rapid Assessment Method (MIRAM). MIRAM provides a rating system that compares a wetlands functional value (including integrity, ecosystem services, and social value) with other wetlands in the state, regardless of ecological type.

B. **Identify a candidate set of wetlands that span the reference gradient in northern Indiana and southern Michigan** (Omernik level 3 ecoregions 55, 56, and 57). We used an objective screening process (remote sensing based metrics and previous ground surveys based on natural heritage methodology) to identify candidate sites that spanned the reference gradient (minimally disturbed to highly degraded). All major wetland types were included. Our remote sensing metrics relied on a combination of landscape condition and stressor metrics relevant to ecological integrity.

C. **Test the effectiveness of the ecological integrity assessment method**on the candidate set of wetland sites and revise the method accordingly. We collected EIA data on 277 sites that comprised the reference gradient, spanning wetland types and conditions. We statistically tested the metrics, including use of a Human Stressor Index, to determine which are most helpful in identifying the range of ecological integrity. We revised the EIA methods and then summarized the ratings for each metric, ecological attribute, and overall set of rank factors. We used an ecological integrity "scorecard" to display the metric ratings and generate an overall **index of ecological integrity**. The scorecard helps interpret the overall status and trends in ecological integrity at a site and across the region.

Section A

We developed our wetland ecosystem condition assessment using the concept of ecological integrity. Building on the related concepts of biological integrity and ecological health, ecological integrity can be defined as "an assessment of the structure, composition, and function of an ecosystem as compared to reference ecosystems operating within the bounds of natural or historic disturbance regimes." To have integrity, an ecosystem should be relatively unimpaired across a range of characteristics and spatial and temporal scales. This broad definition can serve as a guide to developing ecological integrity assessment methods that are distinct from related assessment methods for ecological functions or ecosystem services.

Our multi-metric approach for our Ecological Integrity Assessment (EIA) method is similar to the Index of Biotic Integrity (IBI) for aquatic systems. Our method builds on the work of other rapid assessment methods (especially the Ohio Rapid Assessment Method and California Rapid Assessment Method), and our previous work on standardized methods for assessing ecosystems condition for the natural heritage network along with setting performance standards for wetland mitigation. Critical to our effort was the use of conceptual models that highlight **ecological factors** and attributes for which **metrics** (or **specific indicators**) of integrity are most needed. We defined metrics as values derived from specific measures (e.g., basal area, stand structural class, species diversity) that inform us about the status of an ecological factor or attribute of integrity. For our model, the primary rank factors and major ecological factors were landscape context (landscape, buffer), size, and condition (vegetation, soils, and hydrology). We then selected key metrics that are most responsive, practical, cost-effective and well-tested in measuring the condition of the ecosystem. The conceptual model also provided a structure in which to identify known stressors, or agents of change, that affect these major ecological factors. Together they can help guide management decisions to maintain or restore ecological integrity.

The EIA method also is expected to function across a wide range of wetland ecosystem types. Itallows for various levels of assessment (remote sensing and field based, both rapid and intensive sampling methods), and is structured around the availability of a wide set of indicators and metrics. We provide an overview and demonstration of these methods, with metrics and scoring at multiple scales of assessment, and a scorecard summary of the metrics using an index of ecological integrity (IEI).

Section B

The primary focus of this section was to create a sampling design that would allow us to test the sensitivity of the EIA method to changes in ecological integrity across the full range of wetland types (e.g., bog, rich fen, marsh, wet meadow, wet prairie, swamp, floodplain forest) and conditions (minimally disturbed to degraded). Secondarily, if the sampling design was successful, it could also serve as a screening method for identification of a wetland reference gradient (a set of sites that represent the range of conditions, from minimally disturbed to degraded). Thus, if we successfully create a sampling design that hypothetically spans a range of conditions, and is independently verified, then that design can predict the reference gradient.

Our primary source of potential sites came from the network program databases in Michigan and Indiana. Both programs have identified high-quality (minimally disturbed) "element occurrences" (EOs) of wetland types across the state. These include locations of wetland types with sufficient size and condition to have conservation value. Programs use their own state natural community classifications, which link to the U.S. National Vegetation Classification (USNVC) and NatureServe's Ecological Systems classification. These reference datasets are unique within these states, as they are for many other states where natural heritage programs have gathered data on high-quality examples of native ecosystems.

Until recently, natural heritage programs typically evaluated the condition or ecological integrity of occurrences in the field using best professional judgment and with a minimal amount of quantitative information. The evaluations are summarized using an element occurrence rank (EORANK): A (Excellent), B (Good), C (Fair), D (Poor). Typically, on-site condition was the primary focus when assigning an EORANK. We compiled element occurrences (EOs) for several EPA ecoregions of interest across southern Michigan and northern Indiana. From the available data, we established a site selection process based upon: 1) wetland type and 2) wetland condition. First, we assigned each wetland element occurrence to the macrogroup level of the USNVC (seven macrogroups—Bog & Poor Fen, Rich Fen, Wet Prairie, Wet Shrub, Meadow & Marsh, Coastal Plain Pondshore, Swamp, and Floodplain Forest). This was straightforward based on the cleanly nested crosswalk of state type to macrogroup type. More challenging was the condition stratum, where our best source of information is the EORANK. To increase the standardization of the EORANK, we combined the rank (which emphasizes on-site condition) with a remote-sensing based landscape-context evaluation. This evaluation uses three primary metrics: naturalness of surrounding landscape, land uses within the landscape, and the extent and condition of the buffer immediately surrounding the wetland. When combined, the landscape metrics and natural heritage rank create a "condition stratification rating" for each occurrence (landscape context rating + EORANK rating = condition stratification rating). Minimally disturbed (A and B ranked sites) are often hard to find in this region because of extensive land conversion to agriculture and other land uses. Fortunately, natural heritage databases tend to emphasize identification of those sites.

To ensure sufficient replication of conditions and wetland types, we developed a pool of 280 possible sites (7 macrogroups x 4 conditions x 10 replicates). The natural heritage databases provided most of the occurrences needed to fill this design including many minimally disturbed occurrences. But these databases often lacked degraded occurrences, especially for more common wetland types. We addressed this by having crews send in possible sites based on drive-bys, and then coupled their evaluation with the landscape metrics to see if the site qualified as degraded. Field crews tracked their progress in meeting the sampling design to ensure a representative coverage across the reference gradient. Over time, crews failed to find some sites and other sites were destroyed or had their original reported conditions change considerably. In addition some types (Bog & Poor Fen) are relatively rare and in difficult to access locations, and few degraded examples were available. Our final survey design had 277 sites.

To test the merits of our sampling design, we tallied the number of sites sampled by macrogroup and condition. We successfully maintained a balance across the classification stratum—each of the seven macrogroups had between 30 and 55 survey sites. Thus the predicted wetland type at each site was typically found during the survey. We were less successful in maintaining a balance of sites across the full range of condition, partly because our chosen stratification (screening) method under-predicted the expected number of A condition sites, and over-predicted the number of D condition sites (based on the outcomes of our field assessment of ecological integrity, summarized in Section C). As a result, we reexamined the stratification approach, and proposed a modified version that we recommend for future selection of reference sites.

Thus, our sampling design was sufficiently robust to capture the full range of wetland types and of conditions (even if not strictly balanced). Second, by redesigning our stratification approach, we are confident that we can predict site locations for a reference gradient of wetland types and condition (from minimally disturbed to degraded). Given the widespread availability of both natural heritage data and interpreted remote sensing imagery, we suggest that our methods can be used by studies that need to identify either benchmark (minimally disturbed) sites, or an entire reference gradient. Knowledge of these sites is becoming increasingly important, given increasingly degradation and loss of native ecosystems across many parts of the country.

SECTION C

In this section we tested and applied our ecological integrity assessment method by evaluating 277 wetland sites in the field. Our conceptual model provided a framework to identify metrics for major ecological factors (MEFs), including vegetation, hydrology, soils, size, buffer and landscape. For the rapid assessment (Level 2 or L2), 18 major metrics across all MEFs were initially used, along with an evaluation of stressors to these major attributes. For the intensive assessment (Level 3 or L3) conducted on one-third (88) of the sites, we focused on vegetation measures and metrics. Crews recorded all species and their cover in a 0.1 ha plot. Stem diameters and density for all live and dead tree stems $\geq$ 10 cm dbh were also collected.

All data were entered and managed in an Ecological Observations Database that was specifically designed for the project, yet structured as generically as possible to provide an ongoing database tool for other ecological integrity assessment projects. The database is structured to match field data protocols: General Site Description, L 2 metrics, L 2 stressor checklists, and L3 metrics, including vegetation plot data. Data in 2009 were stored separately for IN and MI. In 2010, several changes were made to the protocol, particularly for stressors checklists, necessitating a slightly different design. The 2010 data from both states were managed in a single database. An Index of Ecological Integrity (IEI), including a scorecard, was used within the database to summarize all metric ratings for L2 assessments. Components of the Floristic Quality Index (FQI) were used to assess integrity for L3. Data were exported from the database in formats suitable for statistical analysis. We created a Human Stressor Index (HSI) based on aggregating stressor scores for Hydrology, Soils, and Buffer. Our primary analyses consisted of screening the metrics based on redundancy among metrics and discriminatory power in relation to the HSI classes. Data are available from NatureServe and from the network programs upon request.

Based on redundancy analysis of metrics, we found that two pairs of metrics had high redundancy (connectivity vs. land use index, and native plant species cover vs. invasive plant species cover). Conversely, among the vegetation metrics, organic matter and increasers had the lowest correlations with other MEFs and to FQI metrics. This suggests that they were not useful metrics for assessing ecological integrity. Based on discriminatory power, soil disturbance and water quality poorly differentiated sites among the HSI classes. Almost all vegetation metrics had low scores in discriminating among HSI classes, which may reflect both lack of discriminatory power to abiotic stressors and responses to other, biotic stressors (e.g., logging, deer browse). The single vegetation metric most responsive to the HSI was vegetation composition. Despite these individual issues with metrics, the overall Index of Ecological Integrity (IEI), the Rank Factors (Landscape Context, Size and Condition), and the Hydrology and Vegetation MEFs were effective in discriminating among HSI classes. Only Soils was not.

Considering the above, we dropped three metrics from our assessment (organic matter, increasers-cover, and water quality). We kept the land use index, so we can test it across a greater variety of land uses across the country. We suggest that native plant species cover and invasives species cover could be treated as two parts of a Native–Invasive Species Index. We also did not drop the Soil Disturbance metric, because we would like to test it on wider range of degraded wetlands, where greater levels of soil disturbance may be expected. But we gave Soils only half the weight of the other two MEFs. Our redesign provides a more equal set of metrics across all MEFs than the original design (especially for vegetation, where the 7 metrics are reduced to 4). Our final recommended list of metrics for ecological integrity assessments of wetland are summarized in the table below:

RANK FACTORS	**MAJOR ECOLOGICAL FACTORS**	**METRICS**
LANDSCAPE CONTEXT	LANDSCAPE	Connectivity
		Land Use Index (optional)
	BUFFER	Buffer Index
SIZE	SIZE	Relative Patch Size (ha) (optional)
		Absolute Patch Size (ha)
CONDITION	VEGETATION	Vegetation Structure
		Regeneration (woody)
		Native Plant Species Cover
		Invasive Exotic Plant Species Cover
		Vegetation Composition
	HYDROLOGY	Water Source
		Hydroperiod
		Hydrologic Connectivity
	SOIL	Physical Patch Types
		Soil Surface Condition

Applying the Final Model

As a final check on the consistency of the method with the best professional judgment methods of earlier natural heritage methods (EORANK), we compared the IEI scores with an independent rescoring of the same sites by natural heritage staff in Michigan, who rate Condition, Size, and Landscape Context, as well as assign an overall EORANK. We found that the Vegetation MEF of the IEI had a very high correlation with the Michigan Condition rating. But other correlations were weaker. We found that the EORANK methods relied more strongly on vegetation, and less on landscape context, hydrology, and soils than the IEI does. We suggest that an overall IEI is the most reliable way to evaluate current conditions of a wetland, including both biodiversity and ecosystem processes.

In conclusion, we demonstrated that our multi-metric EIA method can be effectively used in the field to establish a general index of ecological integrity, in a practical, ecologically meaningful way. Although some of our metrics require greater expertise than others, all attributes have at least two metrics that can be evaluated in a relatively straightforward manner, allowing for wide applicability. The method will have great value for the NatureServe network, contributing to a consistent evaluation of reference sites and the potential for establishing a network of reference standard (minimally disturbed) sites within and across states. Many of these metrics are also in use by other standardized rapid assessment methods (RAMs), including the USA RAM that is part of EPA's National Wetland Condition Assessment. Results here can be used to both refine those methods and provide compatible information on wetland condition across programs. And, in so far as evaluating ecological integrity is a goal within other programs, the EIA method can be a component of those programs, including for inventory, ambient monitoring of wetland condition, and wetland mitigation and restoration.

Section A: The Ecological Integrity Assessment Method

A.1. Introduction

Ecosystem monitoring and assessment programs are critical for resource management, given how ecosystems vary in type, size, landscape settings, and the dramatic losses and degradation that have occurred. These programs are increasingly addressing not just the loss of native ecosystem acres, but also their condition. Data on the ecological condition of ecosystems can be used for ambient monitoring of status and trends, to prioritize sites for conservation or restoration, guide mitigation applications at site and watershed or landscape scales and contribute to land use planning (Fennessy et al. 2007, Faber-Langendoen et al. 2008).

Agencies and organizations are in need of ecological methods that can address this aspect of ecosystems. For example, as part of the National Wetland Condition Assessment in 2011, the Environmental Protection Agency (EPA) carefully designed a comprehensive field survey methodology to assess wetland condition, relying on a reference site approach to establish the criteria for wetland condition (EPA 2011). They were able to draw on a growing body of assessment methods that provide standardized field sampling and reporting methods for assessing ecological condition (e.g., Mack 2001, 2004, Herrick et al. 2005, Pellant et al. 2005, Collins et al. 2006, Fennessy et al. 2007).

There are a number of ways to approach condition assessments, and it is important to clarify the conceptual bases for doing so, in order to ensure that the methods address their intended goals. One important basis on which to assess condition is that of ecological integrity (Andreasen et al. 2001). Building on the related concepts of biological integrity and ecological health, ecological integrity is a broad and useful endpoint for ecological assessment and reporting (Harwell et al. 1999). Ecological integrity can be defined as "an assessment of the structure, composition, and function of an ecosystem as compared to reference ecosystems operating within the bounds of natural or historic disturbance regimes" (adapted from Lindenmayer and Franklin 2002, Young and Sanzone 2002, Parrish et al. 2003). "Integrity" is the quality of being unimpaired, sound, or complete. To have integrity, an ecosystem should be relatively unimpaired across a range of characteristics and spatial and temporal scales (De Leo and Levin 1997). This broad definition can serve as a guide to developing assessment methods, steering us through the related assessment methods for ecological functions or ecosystem services (Dudley et al. 2005, Jacobs et al. 2010).

Ecological integrity concepts are similar to the Index of Biotic Integrity (IBI) concept for aquatic systems. The original IBI interpreted stream integrity from twelve metrics reflecting the health, reproduction, composition and abundance of fish species (Karr and Chu 1999). Each metric was rated by comparing measured values with values expected under relatively unimpaired (reference standard) conditions, and the ratings were aggregated into a total score. Building upon this foundation, others suggested interpreting the integrity of ecosystems by developing suites of indicators or metrics comprising key biological aspects of ecosystems, such as Vegetation IBIs (Mack and Kentula 2010), or more broadly to included, biological, physical and functional attributes of those ecosystems (Harwell et al. 1999, Andreasen et al. 2001, Parrish et al. 2003).

To be effective, the EIA method needs to account for the wide range of ecosystem types (ultimately including terrestrial, freshwater and marine systems), the need for various levels of assessment (remote sensing and field based, both rapid and intensive sampling methods), and the availability of a wide set of indicators. Here we address terrestrial (dryland and wetland) systems.

Critical to this endeavor is the use of conceptual models that highlight ecological attributes for which indicators of integrity are most needed. A conceptual ecological model delineates linkages between key ecosystem attributes and known stressors, or agents of change. It helps identify the ecological attributes we most need to understand regarding the ecological dynamics of the ecosystem, and which we must address when making management decisions to maintain ecological integrity (Noon 2003).

Our goal is to present an overview of our ecological integrity methods, including a) the role of conceptual models and indicators, relying in part on understanding their ranges of natural variability, b) selection of indicators that assess the main ecological attributes and help inform changes that reflect degradation, c) consider indicators at multiple levels of assessment (remote, rapid, intensive), and d) scoring and integrating the indicators in an index of ecological integrity through a scorecard.

A.2. Background

Natural Heritage Methodology

For more than thirty-five years, NatureServe and its predecessor organizations have advanced approaches for documenting the viability and ecological integrity of individual occurrences of species and ecosystems,[1] often referred to as the "elements" of biodiversity (Stein et al. 2000, NatureServe 2002, Brown et al. 2004). Natural heritage methodology often uses the term "element occurrence rank" (EORANK) when referring to the ecological integrity of these ecosystem element occurrences (EOs).[2] Earlier methods relied on fairly qualitative, expert-driven protocols. More recently, the natural heritage methodology has been revised to better reflect an indicator-based approach, one that emphasizes specific indicators to assess the ecological integrity of aquatic, wetland, and dryland ecosystems. Previous publications have provided some of the background (Faber-Langendoen et al. 2008, Tierney et al. 2009, Unnasch et al. 2009); here we provide a major overview to the methods.

Project Organization

The project was funded by the Environmental Protection Agency (EPA). The primary organizations involved in the project are NatureServe, the Michigan Natural Features Inventory (MNFI) and the Indiana Natural Heritage Program (INNHP), with data being made available to Michigan Department of Environmental Quality and others. The development of the Ecological Integrity Assessment (EIA) method, including the conceptual model, occurred over a number of earlier projects, but we provide

---

[1]The Natural Heritage methodology was originally developed by "Natural Heritage" staff of The Nature Conservancy (TNC), many of whom then transferred to NatureServe when it was formed in 2000. Since then, NatureServe staff have worked with the Network of Natural Heritage Programs to maintain and improve the methodology, while continuing to collaborate with TNC staff.

[2] In addition to calling ecological integrity an "element occurrence rank" (or EORANK), Heritage methodology also refers to ecological integrity criteria as "Element Occurrence Ranking Specifications" or EORANKSPECS. Occurrence requirements and mapping guidelines are referred to as "Element Occurrence Specifications" or EOSPECS. We introduced the term "ecological integrity assessment" because it is the more widely used term in conservation biology.

important updates here, including its application across all three levels of assessment (remote, rapid, and intensive). This project was coordinated by NatureServe, whose staff served as the Principal Investigators, with field work contracted to the MNFI and INNHP staff. In addition, staff from EPA provided project oversight, regional input, planning, and technical assistance. Staff from EPA's National Wetland Condition Assessment team and from the Michigan Department of Environmental Quality provided feedback on various aspects of the project.

A.3. PURPOSES OF ECOLOGICAL INTEGRITY ASSESSMENTS

The goal of an ecological integrity assessment is to provide a succinct assessment of the current status of the composition, structure and processes of a particular occurrence of an ecosystem type. These assessments may be done for a number of purposes, including:

- Prioritize occurrences for conservation/management actions (often filtered through site selection criteria).[3] Ratings are helpful both as absolute ratings (best anywhere) and as relative ratings (best of what we have).
- Track status of occurrences over time. After a site is protected and/or put under management, there is a need to know whether the integrity of the occurrence is staying the same or changing. Cost-effective, reliable measures of integrity are needed (Tierney et al. 2009).
- Contribute to information on conservation status. The condition or integrity of occurrences are considered when assigning global, national, and subnational[4] conservation status ranks (GRANKs, NRANKs, and SRANKs), because knowing how many good quality occurrences are on the landscape is an important guide to the overall conservation status or at-risk status of an ecosystem. Conservation status could also be assessed at other scales.
- Prioritize field survey work. Ratings may be used effectively to guide which occurrences should be recorded and mapped (see NatureServe 2002, Section 6, EO Tracking), and to help prioritize occurrences for purposes of conservation planning or action, both locally and rangewide.
- Assess restoration/mitigation efforts. There is an increasing value in using benchmark sites (reference sites) of known integrity to set performance standards for restoration and mitigation—to ensure that wetlands are restored to desired conditions (Faber-Langendoen et al. 2008).
- Inform species population viability ranks. When species populations are closely linked to specific wildlife habitats or ecosystem types, the occurrence rank of the habitat type may serve as a guide for the species viability ratings.

---

[3] Although Element and Element occurrence (EO) ranks help to set conservation priorities, they are not the sole determining factors. The determination of priority occurrences for conservation action will include not only the conservation status of the Element and the likelihood of persistence of the occurrence, but will also include consideration of other factors such as the taxonomic distinctness of the Element; the genetic distinctness of the EO; the co-occurrence of the Element with other Elements of conservation concern at a site; the likelihood that conservation action will be successful; and economic, political, and logistical considerations.

[4] In this document, the term "subnation" will refer to the first order subdivision of a nation (*e.g.*, state, province, district, department).

A.4. THE UNIT OF ASSESSMENT – ECOSYSTEM OBSERVATIONS

Ecosystem Types

The term "ecological integrity assessments" can refer broadly to many aspects of the natural world. Including species (here perhaps the term population viability is more widely used) and other focal resources (Unnasch et al. 2009). Our focus here is on ecosystem types, at multiple scales. On the one hand, we use the term "ecosystem" in a general sense to cover a wide range of scales from macro-scale (such as formations and biomes or broad wetland classes), through mesoscale units (such as macrogroup, group, ecological systems, order, regional wetland classes), to micro-scale units (alliances, associations, natural community types etc) (see Faber-Langendoen et al. 2008). On the other hand, we use ecosystem in a particular, and pragmatic sense, as a spatial entity bounded on the landscape by a specific set of biotic and abiotic structural, compositional and ecological criteria, at scales that address the most common concerns of conservation and management.

We refer to the specific place where an ecosystem type is found as an "ecological observation." Many other terms are used, such as "assessment area," "sample point", "ecological site," "field site," "occurrence," or "stand." The term "observation" is sometimes used as a generic, flexible term applied to any kind of place or unit where an ecosystem is identified and described (Stevens and Jensen 2007), and is increasingly used a term for all species or ecosystem field records (Lapp et al. 2011). The EIA method focuses on understanding the structure, composition, and processes that govern the wide variety of ecosystem types at particular sites where an ecosystem is found. Ecosystem classifications are important tools in assessing the ecological integrity of these observations. They help ecologists to better cope with natural variability within and among types so that differences between observations with good integrity and poor integrity can be more clearly recognized. Classifications are also important in establishing "ecological equivalency," for example, in providing guidance on how an impacted salt marsh can be restored to a salt marsh with improved integrity.

Points, Polygons, and Patches

Assessments of ecosystem condition can be based on observations defined as points, polygons, or patches.

A point based approach in which a fixed area is sampled around a point offers some advantages (Fennessy et al. 2007, Stevens and Jensen 2007):

- simplicity in terms of sampling design
- no mapped boundary of ecosystem type is required for assessment unit
- limited practical difficulties in the field of assessing the entire area, as the area is typically relatively small (0.5–2 ha); long-term ambient monitoring programs often use a point-based approach because of these advantages

A polygon approach, in which a specific ecosystem area is delineated (using vector or raster methods), offers some advantages:

- Mapping boundaries facilitate whole ecosystem and landscape interpretations
- Decision-makers and managers are often more interested in "stands" or "occurrences," rather than points

- Programs that maintain mapped occurrences of ecosystem types are most interested in the status and trends of those occurrences

Pixel- (or raster-) based imagery, such as from satellites, are perhaps intermediate between points and polygons. Pixels are often smoothed into larger "patches," these patches can be assigned to ecosystem types, and analyses can be performed on these patches, or on a series of patches. These series may be created as clusters (e.g., using separation distances between patches, comparable to cluster polygons) or as "bounded patches," where a larger landscape or watershed boundary is used, and all patches of the same ecosystem type within that boundary are included as part of the assessment area. The "bounded patch" approach is currently being used by NatureServe to conduct ecological integrity assessments in western U.S. ecoregions (NatureServe 2011 in prep).

A raster- or vector-based polygon approach is widely used for inventory purposes by many agencies. Many programs that are part of NatureServe network routinely classify and map ecosystem occurrences, and maintain extensive information on the structure, composition and stressors to those occurrences. While not without its complications, standard mapping guidelines can be applied and line work can be provided digitally to crews, who can then adjust them as needed directly on aerial photo or digitally on field recorders.

There are also regulatory advantages to polygons, as noted by Fennessy et al. (2007) for wetlands "the basic "currency" in Clean Water Act Section 401/404 regulation of wetlands is something called a "wetland" and this is also the common understanding: a "wetland" is a definable piece of real estate that can be mapped and walked around. There are substantial pragmatic and legal considerations in developing a condition assessment protocol that cannot assess "wetlands."

Where the occurrence at a site is the focus, then a sampling design could still vary as follows:

- conduct an assessment survey of the entire area of the occurrence, e.g., a rapid qualitative assessment;
- conduct an assessment survey of a typical subarea(s) of the occurrence, preferably of uniform condition, or
- collect a series of plots, placed either in representative or unbiased locations, throughout the entire area or subarea occurrence

In all three cases, the intent is to assess the ecological integrity of a particular wetland occurrence. The occurrence may, in fact, be defined by the combination of ecosystem type and level of integrity. Thus a minimally disturbed wetland type can be mapped and assessed separately from a degraded example.

Watersheds & Landscapes

The condition of entire watersheds or ecoregions can also be assessed. Ecoregional status assessments or watershed profiles are ways of characterizing the entire landscape area. We do not apply the term "ecological integrity assessment" to these approaches, as our definition of ecological integrity is *"an assessment of the structure, composition, and function of an* ***ecosystem as compared to reference or benchmark ecosystems*** *operating within the bounds of natural or historic disturbance regimes."* Although ecoregional units can be viewed as "landscape ecosystems (Bailey 1996)," these units more often are viewed as landscape or watershed types within and across which ecosystems are found (Forman 1995); i.e., the ecosystems provide a "bottoms-up" approach where assessments of component ecosystems contribute to an overall rating for the landscape or watershed. Nonetheless, assessing the condition of the landscape area can provide important information on the ecological integrity of the ecosystems within those landscapes.

A.5. Range of Natural Variability and Ecological Integrity

The Range of Natural Variability Concept

Species and ecosystems all evolve within dynamic environments; and naturally exhibit some range of natural variability in their attributes over time and space. For example, the age and species composition of any forest canopy naturally vary over time and from one stand to the next; and any forest naturally experiences varying frequencies and intensities of disturbance from fire, drought, wind damage, or flooding. Similarly, coastal salt marshes naturally experience varying frequencies and intensities of nutrient and sediment inputs, tides, wave action, and storms. Within the limits of this range, further, the variation may be either patterned (e.g., cyclical) or random; and may play out over scales of time from hours and days to decades and centuries.

The natural variation in both space and time is thus essential to shaping ecosystems. Consequently, the natural range of variability depends on specifying the time frame. For purposes of assessment projects, where the horizon is usually 30 to 100 years, we normally treat the natural variability in each key attribute of a system as occurring within stable limits. However, there may be situations in which this is not appropriate.

Resource managers often use the concept of a range of natural variability (RNV) (e.g. Landres et al. 1999, Oliver et al. 2007). Information on RNV provides important clues on the long term driving variables and disturbances that shape ecosystems, the flux and succession of species, and the relative role of humans in shaping the systems. Understanding RNV is important for placing interpretations of ecological patterns in their historical setting. However, what is 'natural' can be difficult to define, given limited knowledge of ecosystems, the extent of past human activity, and the likely effects of ongoing and future climate change. Scientific knowledge of most ecosystems has a relatively short history, as does the preserved record of most environmental regimes (fires, floods, etc.). The variation in ecological dynamics that we observe within years or decades can be part of much larger trends or cycles spanning centuries or millennia. For these reasons, others prefer the term "historic range of variability" (Egan and Howell 2005).

No ecosystem, natural community, or species is ever static when viewed on larger scales of time. Human activity has thoroughly transformed many places throughout the world, and no place is free of human impacts. Much as a changing climate throughout the Holocene (past 12,000 years) brought about changes in many of aspects of ecosystems, and resulted in many patterns of species composition we see today, so too have certain human activities shaped ecosystems. Humans have brought about large-scale and long-term changes in ecosystems even far from our farms and cities, for example through hunting and selective tree removal, releasing non-native species, setting fires, and diverting streams.

In many instances where the rate and magnitude of human-induced change may be limited, we can safely subsume their effects within a practical 'natural' range of variation. That is, we can assume that their effects have had only a limited impact on the evolutionary environment of biodiversity. At the same time, we can often detect human effects that cause rapid and substantial ecological change. And we can do so not only in recent, better documented times but in the more distant past, for example from records of ancient land clearing for corn production, desert stream diversions, or the draining of arable swamplands. When we can detect such more significant human effects, we need to presume them to be outside of some practical, ecologically functional range of natural variability (i.e., likely resulting in local extinctions and other biodiversity impacts).

Ecological Integrity and Range of Natural Variability

Given these challenges, it is important to emphasize what can and cannot be achieved by using RNV as a component of ecological integrity (Higgs and Hobbs 2010). First, it is the knowledge of natural variability that informs our goals and evaluations of current conditions, but this knowledge does not a priori constrain how we state desired conditions (clarity in goals). Second, to suggest that we can simply take over the management of natural ecosystems without understanding RNV is to invite failures in these complex systems (restraint and respect). Finally, the purpose of understanding RSV is not to lock us in the past, but to ensure that we connect the historical ecological patterns and processes to the present and future (historical fidelity). Fourth, understanding RNV will ensure that we can anticipate change and emphasize resilience in the face of future changes (Higgs 2003, Higgs and Hobbs 2010). In this way, NRV as a component of ecological integrity takes us beyond a simple interpretation of what is natural to engaging us to think through how our actions and goals can maintain natural ecosystems. Finally, NRV need not be interpreted solely from the historical record, but from benchmark sites currently present on the landscape.

Because it can be difficult to define what is natural, alternative terms have been suggested, including "acceptable range of variation" (Parrish et al. 2003). The key point is that direct knowledge of the range of natural variability, is but one source of information for developing proposed ratings of ecological integrity. Other sources of information include ecological models, expert knowledge, and comparisons to a reference gradient or reference standard of the same or similar ecosystems (Parrish et al. 2003, Stoddard et al. 2006). Even where present-day reference standard sites may be hard to identify based on minimally disturbed criteria, one may still be able to make reasonable estimates based on historic data or inferred species-habitat relationships (Brewer and Menzel 2009). Particularly where such examples have been affected by human impacts of varying types and magnitudes, comparisons can be especially informative about where the limits may lie beyond which the persistence of the ecosystem may be at risk. Even where present-day reference standard sites may be hard to identify based on minimally disturbed criteria, one may still be able to make reasonable estimates based on historic data or inferred species-habitat relationships (Brewer and Menzel 2009).Thus the ranges of ranges of variability specified in the indicators are ranges relevant to the hypothesized levels of ecological integrity, and our understanding of those ranges will change over time.

Thus, both our understanding of ecological integrity and what is natural change over time. Too often the characterization of integrity is treated as a static linear function, not unlike the model shown in Fig. A1. In the short term, these models can be helpful. But they can be misleading with respect to both the ongoing natural, historical processes that shape ecosystems and the human interactions with those systems. Simplistic views of "natural" as referring only to "pristine conditions" is not tenable, given the long interactions between humans and the environment. But simply collapsing human activity (culture) into an extension of natural processes is also too simplistic. It may be helpful to expand our view by considering how ecology and human culture are "knitted together over time;" that is, both culture and ecology have histories, and consideration of current ecological integrity reflects both histories, without suggesting that they are one and the same (Fig. A2, Higgs 2003). What is critical is to ground our ideas of ecological integrity in the knowledge gained from current reference sites; thereby spanning our cultural perspective on integrity with known ecosystem sites in the present, as informed by the past.

FIGURE A1. Simple schematic showing how ecosystem structure and function may recover over time to either the more original (historical, natural) system or some altered form.

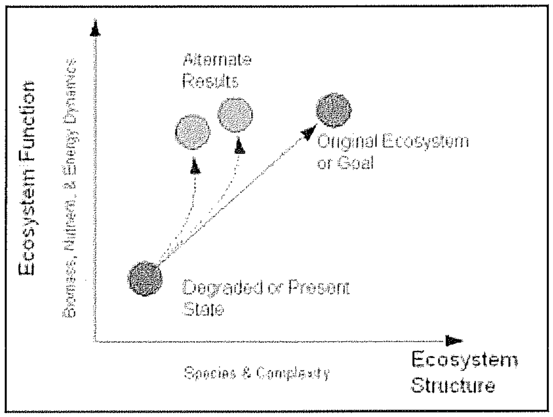
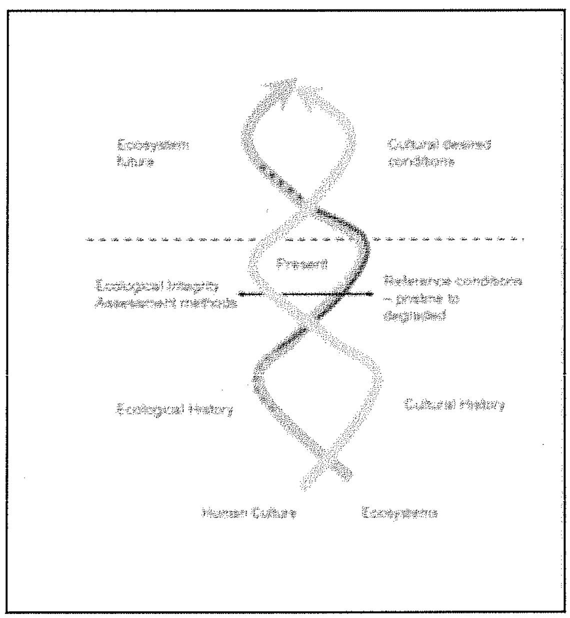


FIGURE A2. A model of ecosystem change showing how ecosystems and culture are interrelated through time. Red/orange color highlights the increasing changes affecting ecosystems, and the uncertainty of those changes into the future (Adapted from Higgs 2003, Fig. 6.2).



Ecological Integrity and Climate Change

Global climate change, a further consequence of human activity, is bringing about changes in regional and local climate. Every place on Earth now faces changes in the magnitude, timing, frequency, and duration of atmosphere-driven conditions—from changes in seasonal temperatures and weather patterns to changes in the temperature and pH of our oceans—many potentially outside the range of historic variation. The ecosystems of tomorrow in every region potentially will experience ranges of variation in atmosphere-driven conditions far different than have the ecosystems in these regions even in the recent past.

Understanding the natural or historic range of variability has immense value in improving our understanding of ecosystem responses to environmental changes and setting management goals (e.g., Swetnam et al. 1999). However, as it now apparent from our discussion above regarding ecological integrity, it may no longer be sufficient to assume that establishment of historical conditions will assure that ecosystems will persist into the future, in the face of novel anthropogenic stressors, such as pollution, habitat fragmentation, land-use changes, invasive species, altered natural disturbance regimes and climate change (Millar et al. 2007). Thus measures of ecological integrity need to account for the ability of ecosystems to "adapt" to changes, as climate and environments shift. These shifts may create environments that are outside any known range of natural variation ("novel climates" of Williams et al. 2007). This does not make the past and current states irrelevant; rather, as Millar et al. (2007) note: "Historical ecology becomes ever more important for informing us about environmental dynamics and ecosystem response to change."

A.6. Ecological Integrity Assessment Method

Our method is based on the following set of key steps:

1. determine the purpose of the assessment
2. develop a general conceptual model for wetlands, adapted, as needed, for various ecosystem types
3. rely on indicators of ecological attributes that span the major structural, compositional and ecological processes of the system
4. select indicators across three levels of assessment—(i) remote sensing, (ii) rapid ground-based, and (iii) intensive ground-based metrics
5. scale the thresholds or assessment points of the indicators based, in part on ranges of natural and historic variability, ecological models, benchmark or reference sites
6. summarize indicators using ratings and integrate into an overall index of ecological integrity

Purpose of the Assessment

In the section above, we noted some broad purposes of ecological integrity assessments, such as, 1) prioritizing observations for conservation/management actions, 2) tracking status of observations over time, 3) assess management actions, 4) contribute to range-wide conservation status of ecosystems, and 5) provide performance standards for mitigation and restoration. These general goals need to be further refined to make sure the assessment is structured to address the needs. The geographic scale of the assessment also has an impact e.g., national trends monitoring, regional landscape assessment, local landscape assessment, local site management and monitoring. These varied purposes require that the approach to EIAs be flexible, while retaining a consistent core of ecological attributes. To do so, we develop a general conceptual model, and then suggest ways in which its application can be applied in a flexible manner.

Conceptual Model for Terrestrial Ecosystems

Identifying the ecological attributes that need to be assessed involves building a conceptual ecological model of ecological integrity. This model rests on the knowledge of the system, its setting, and similar or associated systems. The result is a set of hypotheses about how the system functions, its defining characteristics and dynamics, and critical environmental conditions and disturbance regimes that may act as drivers of these characteristics and dynamics. These hypotheses both guide management and monitoring, and highlight gaps in knowledge that require additional investigations (Unnasch et al. 2009).

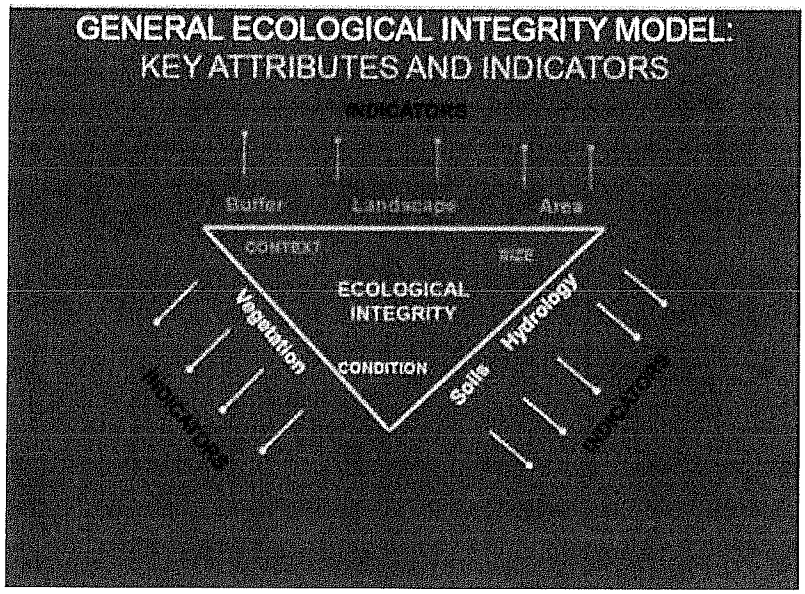
We use a conceptual ecological model that provides a general set of ecological factors common to all terrestrial (wetland and upland) systems, and then encourage identification of individual key ecological attributes for individual system types. The model also provides a means to assess stressors or agents of change to the ecological factors (Noon 2003). The terms for the model come from a variety of models available in the literature (Table A1) and that of Faber-Langendoen et al. (2006).

TABLE A1. Comparison of terminology among various agencies and organizations for ecological integrity / assessments (modified from Faber-Langendoen et al. 2006). Overarching goals and objectives are defined variously by each group. TNC = The Nature Conservancy, EPA = Environmental Protection Agency, NPS = National Park Service.

NatureServe	TNC	EPA	NPS Vital Signs
Rank Factor		GoalObjective	
Major Ecological Factor		Essential Ecological Attribute (EEA)	Level 1 Category
		EEA subcategory	Level 2 Category
Key Ecological Attribute	Key Ecological Attribute	Ecological Endpoint	Level 3 Category (Vital Sign)
Indicator/Metric	Indicator	Measure	Measure / Metric
Comer et al. (2003), Faber-Langendoen et al. 2008)	Parrish et al. 2003	Harwell et al 1999, Young and Sanzone 2002	Fancy et al. 2009

The major components of the model include three primary rank factors (landscape context, size, and (on-site) condition, subdivided into 6 major ecological factors: landscape, buffer, size, vegetation, hydrology, and soils. Together these are the components that capture the structure, composition and processes of a system (Fig. A3). Other major attributes, such as birds, amphibians, and macroinvertebrates, can also be assessed where resources, time and field sampling design permit. The model is fairly intuitive, but a key component is that, to describe how a system "works," one must include both the "inner workings" (condition) and the "outer workings" (landscape context). Assessing size of ecosystems helps to characterize patterns of diversity, area-dependent species, and resistance to stressors. Conservation of such characteristics and processes will contribute not only to current ecological integrity but to the resilience of the ecosystem in the face of climate change and other global causes of stress.

FIGURE A3. Example of conceptual model for ecological integrity assessments of terrestrial ecosystems. The core ecological attributes of ecological integrity are shown for wetland and uplands. The model can be expanded to include additional measures of biotic integrity, such as birds, amphibians, macroinvertebrates, etc.



Rank Factors and Major Ecological Factors

An **ecological factor**—sometimes referred to as a major ecological attribute (Faber-Langendoen et al. (2008))—of an ecosystem is common to all ecosystems and encompasses a broad set of related ecological attributes. Thus it serves to ensure that assessment methods will not neglect major structural, compositional and ecological processes of the system. These ecological factors can also be used as broad indicators in their own right, and may be assessed qualitatively, using narrative approaches, in very rapid assessments.

For our wetlands model , we use three primary rank factors—Landscape Context, Size, and Condition, as is typical of a number of major assessments (NatureServe 2002, Parkes et al. 2003, Oliver et al. 2007). Together these primary rank factors provide the orientation for more specific selection of major ecological factors, key ecological attributes and indicators.

The primary and major ecological rank factors are:

- "Landscape Context" refers both to the spatial structure (spatial patterning and connectivity) of the surrounding **landscape** and the immediate **buffer** within which the ecosystem occurs; and to critical processes and environmental features operating at landscape scales that affect the system. Examples of landscape structure include attributes of fragmentation, patchiness, and proximity or connectivity among habitats. Examples of landscape processes include the interaction of matter, energy, and disturbance regimes between a focal ecological system and surrounding systems.

- "Size" refers to attributes related to the absolute or relative **size** of the ecosystems, measured as geographic extent, or area, or the size of the natural patch within which a system is found.
- "Condition" refers to critical biotic and abiotic structure, composition, and processes, and their dynamics, typically subdivided into **vegetation**, **hydrology** and **soils**. Examples include species composition, hydroperiod, soil chemistry, grazing intensities, and vegetation structural stages arising from natural disturbances.

Although the model presumes that all systems share the primary and major ecological factors, the weights given to them may vary. For example, hydrology can be given much more weight in wetland systems. Size may be of minor importance in spatially patchy wetland systems. Many standard wetland assessments emphasize these ecological factors—buffer, vegetation, hydrology, and soils (Fennessy et al. 2007). Consistent recognition of these ecological factors also helps interpret the information coming from a variety of assessment programs, particularly as these vary in the role given to interpreting the biotic versus abiotic assessment of condition. Programs that use Indices of Biotic Integrity are relying on the biota or vegetation as the primary or sole measure of condition. Natural heritage programs often use vegetation as the primary means of assessing condition. For landscape context, many programs measure both the buffer around a point or polygon and the larger landscape (e.g., Jacobs et al. 2010).

Size provides some challenges to an integrity assessment. Some ecosystems types vary widely in size for entirely natural reasons (e.g., a forest type may have very large occurrences on rolling landscapes, and be restricted to small occurrences on north slopes or ravines in other landscapes). For other types, size may be relatively unimportant because they are always small (seepage fens, talus slopes). Nonetheless size can be an important aspect of integrity. For some types, diversity of animals or plants may be higher in larger occurrences than in small occurrences that are otherwise similar. For occurrences that occur in mosaics, the larger occurrences often have more habitats. Larger wetlands are also more resilient to hydrologic stressors, since they buffer their own interior portions to some extent. Studies have also shown that wetland size is a strong predictor of wetland condition, probably as a function of landscape fragmentation (Fennessy et al. 2009). Thus size can serve as a readily measured proxy for the interdependent assemblage of plants and animals, particularly in the context of more rapid assessments, where limited measurements can be taken. Finally, we can separately assess absolute size from relative size (the size a patch would have under natural conditions versus stressed condition (e.g., a small wetland that has been partially filled in or drained). For all these reasons, we retain size as a high level indicator, but keep it separate from other indicators, so that its contribution in the overall rating of integrity is clear.

Recognized stressors (threats) to an ecosystem also provide crucial information for the identification of key ecological attributes. Stressors to the system include human activities, structures, or institutions—or consequences of these. They alter one or several key ecological attributes beyond their acceptable ranges of variation. Consequently, knowledge of how specific human actions cause harm to an ecosystem can provide insight into the resource's key ecological attributes, and vice versa. Again, to simplify these steps and to ensure consistency across models, we suggest that stressors should be tracked across all ecological factors, but can be fine-tuned based on knowledge of key ecological attributes.

Key Ecological Attributes and Indicators

A **key ecological attribute** (KEA) of an ecosystem is one that is critical to a particular aspect of the ecosystems persistence in the face of both natural and human-caused disturbance, and alterations to that attribute beyond some critical range of variation will lead to the degradation or loss of that ecosystem. It is much more directly measurable than are ecological factors and more amenable to an indicator-based approach.

Metrics are sometime referred to as **indicators,** and here we use the term in the sense of a fine-grained indicator. For clarity, we distinguish "metrics" from both "measures" and "general indicators." **Measures** are those values that are collected directly in the field (e.g., diameter of tree at breast height, species percent cover) and metrics are values derived from specific measures (e.g., basal area, stand structural class, species diversity) that inform us about the status of an ecological attribute of integrity. Coarse woody debris is a general indicator, whereas volume of coarse woody debris and biomass of coarse woody debris are two closely related metrics for that indicator. For a given system, a single metric is typically used for an indicator, but when comparing across systems, we may have different metrics for the same indicator. Standardization of metrics is certainly helpful, but more important are agreement on the kinds of general indicators that are used to characterize ecological integrity, at least among related ecosystem types.

To be a **metric (or specific indicator)** requires that it be informative about alterations to the attribute that may lead to the degradation or loss of the ecosystem.

The metric may be either:

- A specific, measurable characteristic of the major or key ecological attribute (e.g., percent cover of native species, coarse woody debris, calcium:aluminum ratio of soils, hydroperiod).
- A collection of such characteristics combined into a "multi-metric" index, such as a forest structural stage index that integrates measures of tree size across all stems, or a buffer index that combines buffer length, width and condition, or
- A measurable effect of the key ecological attribute, such as a ratio of the frequencies of two common taxa of aquatic insects (the indicator) that varies with changes in average Nitrate concentration (the key attribute) in a stream.

Metrics are selected to meet ecological, technical and management needs (Tierney et al. 2009, Unnasch et al. 2009, Fancy 2009).

<u>Ecological – Statistic Criteria</u>

1. Specific (redundancy): unambiguously associated with the key or ecological factor of concern and not significantly affected by other factors

2 . Sensitive (discriminatory power): able to detect changes that matter to the persistence of the ecosystem

3. Comprehensive (range): able to detect changes across the entire potential range of variation in the key ecological attribute, from best to worst condition

Technical Criteria

4. Measurable: measurable by some procedure that produces reliable, repeatable, accurate information

5. Technically feasible: amenable to implementation with existing technologies without great conceptual or technological innovation

Management Criteria

6. Timely: able to detect change in the key ecological attribute quickly enough that project managers can make timely decisions on conservation actions

7. Cost-effective: able to provide more or better information per unit cost than the alternatives

8. Partner-based: compatible with the practices of key partner institutions in the conservation effort, or based on measurements they can or already do collect

9. Legal mandates: addresses legal requirements for the program

It is rarely possible to identify a single indicator that meets all eight criteria for an individual key ecological attribute, but ensuring relevance across the broader categories is important. Managers may need to put several indicators together to obtain a more reliable or more complete picture of the system. For example, indicators taken from both field surveys and analyses of aerial photographs may provide complementary and more reliable assessments of forest tree composition, or of characteristics of the buffer around a wetland, than either indicator can on its own.

A variety of statistical methods are available to help assess the statistical rigor of metrics, applicable to both rapid and intensive metrics. The most readily assessed criteria include comprehensive range, discriminatory power or responsiveness, and redundancy (Blocksum et al. 2002, Klemm et al. 2003, Jacobs et al. 2010).

As an example of a good indicator is that of, "Relative Total Cover of Native Plant Species." It is measured by estimating total cover of all exotic species subtracted from total cover of all vegetation and divided by 100" (Table A2). It is specific, measurable, fairly sensitive (though separating invasive exotics from other exotics may increase its sensitivity), timely (again distinguishing invasives from other exotics might increase its timeliness), technically feasible (the "exotics" category is generally well defined in botanical manuals), cost-effective and partner-based (many organizations and agencies are interested in the presence and abundance of exotics). It may be one of several specific indicators for a Vegetation Composition KEA. Similarly, "Woody Regeneration" (such as tree seedling and sapling regeneration in forests) is another specific indicator (metric) for Vegetation Structure (KEA). Together these are indicators for overall Vegetation (MEF). Measurement of indicators is described in protocols that ensure consistent and repeatable measurements. By contrast, leaf and other small woody cover in systems is much harder to specify in terms of natural variability relevant to ecological integrity, particularly in rapid assessments. It may be too variability within a site over the course of a season.

The identification of key ecological attributes and specific indicators for each ecosystem or group of related ecosystems is an iterative process. It may require that KEAs and indicators be identified for each ecosystem type, or they may be applicable across many types. A review of the level of ecological classification may be needed to ensure that KEAs and indicators are ecologically meaningful, applicable to issues of resource management and cost-effective.

TABLE A2. Example of a Vegetation indicator (metric)

RANK FACTOR:	CONDITION
MAJOR ECOLOGICAL FACTOR	VEGETATION
KEY ECOLOGICAL ATTRIBUTE	Vegetation Composition
Metric:	Relative Total Cover of Native Plant Species
Definition:	Percent cover of the plant species that are native, relative to total cover (sum by species)
Metric Ratings	Metric Criteria
A = Excellent	>99% cover of native plant species
B = Good	95–99% cover of native plant species
C = Fair	80–94% cover of native plant species
D = Poor	50–79% cover of native plant species
E = Very Poor	<50% cover of native plant species

To assist with this process, and ensure a level of standardization across models, we suggest that any model should include key ecological attributes and indicators under each of the primary factors. The final list should focus attention on those potential key ecological attributes that are the most defining, most critical or pivotal to the persistence of the ecosystem and its natural internal dynamics. If an attribute or indicator does not appear to be responsive to stressor changes or appears to be unrelated to these major attributes, it is a signal that it is not critical to ensuring the persistence of the ecosystem.

Integrity Metrics and Stressor Metrics

The primary emphasis of the indicators is on measuring a relevant attribute of the ecosystem itself that is clearly related to known ranges of natural variability and that are responding to stressors. We refer to these as "**integrity metrics**" or indicators. We can also measure the stressors themselves, but information from these metrics provides only an indirect measure of the status of the system—we will need to infer that changes in the stressor correspond to changes in the integrity of the system. We refer to these as "**stressor metrics.**" We provide a catalogue of possible stressors at a site (stressor checklists) to guide interpretation and possible correlations between ecological integrity and stressors.

We prefer to use integrity metrics separate from stressors, in order to independently assess the effects of stressors on integrity, but occasionally a stressor metric is substituted for an integrity metric when measuring integrity is challenging or not cost-effective. For example, a "Land Use Index" indicator is a stressor metric that characterizes the level of stress produced by land uses in the surrounding landscape, rather than characterizing the integrity of the ecosystems in the surrounding landscape. The basic goal is an accurate, cost effective estimate of integrity, rather than concern to keep the model pure.

Thresholds in Ranges of Ecological Integrity

The EIA method posits that some degree of thresholds exist in the range of potential variability for each key ecological attribute. These are thresholds, outside of which managers should anticipate—or sometimes may already observe—signs of unacceptable change or degradation to the ecosystem (Mitchell et al. 2011). Ecologists typically cannot estimate specific probabilities of persistence for communities and ecological systems, as can be done for species populations. Instead, we recognize that unacceptable alteration will involve severe degradation of an attribute or the entire system, leading to its transformation into some other kind of system altogether (e.g., the stream flow stops, leaving a dry stream bed; a grassland becomes a woodland in the absence of fire). Such a transformation might begin with the loss of only a few highly sensitive species, although it could increasingly affect the more common and less specialized as well.

The EIA method requires identifying the critical ranges of variability for each indicator used to keep track of each key attribute, for each resource. Of most concern are stressors that cause certain ecological attributes to vary beyond certain threshold values. Critical ("hard") thresholds occur when a given point along a continuum of change in an ecological attributes leads to an alternative state of an ecosystem (*sensu* Holling 1973). Hard thresholds mark specific conditions beyond which ecosystems change irreversibly; soft thresholds track a range of conditions over which ecosystems are changing, with varying degrees of reversibility. An example of hard thresholds is when certain levels of phosphorus loading in streams may cause a series of cascading effects. With soft thresholds, an alternative state may emerge gradually ("soft" thresholds) as the attribute changes (Mitchell et al. 2011). For example, declines in species richness in response to degradation may show a continuous decline, and setting a threshold is more akin to setting benchmarks along a continuum.

When a key ecological attribute crosses either a hard or soft critical threshold, the resource itself may not experience either immediate or abrupt change. The resource may initially only lose its capacity to resist change triggered by new disturbances and/or its capacity to recover following a new disturbance. Once a resource suffers such a loss of resistance or resilience, however, it may take only a slight additional change to trigger further alteration away from its acceptable range of variation. For example, the suppression of fire in an aspen woodland for more than a few decades could leave it vulnerable to the arrival of seeds from other nearby communities, that could lead to the replacement of the dominant tree cover by Douglas fir and other conifers that promote changes in soils and ground-cover vegetation that attract different fauna that further transform community dynamics, and so forth (Unnasch et al. 2009).

The changes that ensue when a key ecological attribute passes some critical threshold may take considerable time to play out, particularly in systems with very long-lived species. Nevertheless, once set in motion, such chains of consequences may be difficult to reverse. The alteration of one or more key ecological attributes beyond their acceptable ranges of variation can reach a further threshold, beyond which the focal resource will almost certainly fail unless the situation is quickly reversed. Particularly worrisome are thresholds of ecological collapse, failure which could mean potentially irreversible transformation into—or replacement by—some other kind of community or system (the "thresholds of imminent loss" of Unnasch et al. 2009).

The ease of crossing threshold may be different from a degrading perspective than aggrading perspective. For example, having crossed the C/D threshold for phosphorus loading, it may be very difficult to restore back to a C. Or having lost a top predator or grazer from a system, it may be difficult to reestablish it. Or once the density of shrubs in a pine savanna increases to a certain level, it may become very hard to reintroduce fires.

Estimating the range of natural variability and assessing thresholds for each indicator answers the crucial questions, how much alteration of a key ecological attribute is too much? Managing ecosystems based on NRV in turn does not mean managing for all the variation that the resource might experience under undisturbed conditions. Instead, it means managing only for an envelope of conditions that together are "sufficient" for resource persistence, function, and for achieving related management goals.

Estimating the range of natural variability for every indicator may be a challenge. It requires some knowledge of the historic range of variability for all key ecological attributes and their indicators. Or it could mean establishing range of variability from a current set of reference sites judged to be in various states of ecological integrity (these can in turn be linked to historic interpretations of NRV at those sites. Fortunately, even initial approximations about the acceptable range of natural variability for an indicator provide hypotheses on which both to begin management and to begin research to improve the initial estimates.

Reference Condition as a Guide to Indicator Rating

A key method for establishing ecological integrity ratings is based on reference sites. We can refer to the full range of reference sites as the **reference gradient** (also referred to as **reference set,** or **reference network**); that is *"the gradient of ecosystem condition across a region varying from least disturbed (reference standard) to highly impaired.* The set of reference sites represent the range of variability that occurs among stands of a wetland type as a result of both natural processes (e.g., succession, channel migration, fire, erosion, and sedimentation) and anthropogenic alteration (e.g., grazing, timber harvest, and clearing) (Klimas et al. 2006).[5] There has been much discussion on whether **reference standard** should be based on "**minimally disturbed wetlands**," i.e., the subset of the gradient of reference wetlands that exhibit metric ratings for the type at a level that is characteristic of the historically and/or currently minimally disturbed wetland sites in the landscapes (Klimas et al. 2006, Stoddard et al. 2006). Using the "minimally disturbed" approach, the reference standard sites would typically have, or be able to attain a high ecological integrity rating for all or most metrics.[6] The geographic area from which reference wetlands are selected is sometimes referred to as the **reference domain** (Smith et al. 1995). The reference domain may include all (ideally), or part (e.g. within an ecoregion), of the full range extent of a type. Where few sites exist today that are minimally disturbed, reasonable estimates can still be made based on historic data or inferred species-habitat relationships (Brewer and Menzel 2009).

---

[5] As Sutula et al. (2006) state, "one important element of metric development is definition of the standard of comparison that defines the highest and lowest levels of potential or expected wetland condition. This standard of comparison is commonly referred to as a reference; however, the concept of reference is more accurately defined as a range of conditions that can be correlated with a known set of stressors. The highest point on this reference continuum is then termed reference standard condition. The collection of sites or theoretical states that represent a gradient in conditions is referred to as the reference network. To the extent possible, the reference conditions should be represented by actual wetlands."

[6] When choosing a reference standard, one needs to choose whether such a standard represents the Minimally Disturbed Condition (MDC) or Least Disturbed Condition (LDC), or a combination of the two, based on best attainable condition (BAC). Huggins and Dzialowski (2005) note that MDC and LDC set the high and low end of what could be considered reference standard condition. They go on to say that "these two definitions can be used to help define the Best Achievable Conditions (BAC's), which are conditions that are equivalent to LDC's where the best possible management practices are in use. The MDC's and LDC's set the upper and lower limits of the BAC's. Using the population distribution of measures of biological condition associated with a reference population might provide some insights regarding the potential relationship between the MDC and LDC for a particular region."

Reference gradients serve several purposes. First, they are the source of information on what constitutes a characteristic and sustainable level of integrity across the suite of ecological attributes selected for a type (i.e., by visiting ecosystems that range from excellent to degraded in integrity we can document the characteristics of each level of condition). Second, reference gradients establish the range and variability of conditions exhibited by assessment metrics and they provide the data necessary for calibrating assessment variables and models (i.e., they help guide the interpretation of wetland status and trends assessments). Finally, they provide a concrete physical representation of wetland ecosystems that can be observed and re-measured as needed (Smith et al. 1995, Klimas et al. 2006).

Indicator Ratings

With a well-chosen set of metrics selected to track changes in the major and key ecological attributes, based on the natural range of variability, with thresholds established for each metrics, we can structure the rating system to guide our interpretation of ecological integrity, using simple scorecard grades:

> **Excellent:** The indicator lies well within its range of natural variability.
>
> **Good:** The indicator lies within but is near to its range of natural variability.
>
> **Fair** The indicator lies outside its range of natural variability, but not outside its threshold of ecological collapse.
>
> **Poor:** The indicator lies near to well outside its threshold of ecological collapse.

Indicators with higher levels of integrity would generally be rated "A", "B", or "C" (from "excellent to at least "fair" integrity), and those with substantial degradation are rated "D" ("poor" integrity) (see Table 1).

Indicators at Multiple Scales (Level 1 to Level 3)

Overview of the Three Levels

The selection of metrics to assess ecological integrity can be executed at three levels of intensity depending on the purpose and design of the data collection effort (Brooks et al. 2004, Tiner 2004, US EPA 2006). This "3-level approach" to assessments, summarized in Table A3, allows the flexibility to develop data for many sites that cannot readily be visited or intensively studied, permits more widespread assessment, while still allowing for detailed monitoring data at selected sites.

TABLE A3. Summary of 3-level approach to conducting ecological integrity assessments (adapted from Brooks et al. 2004, USEPA 2006).

Level 1 – Remote Assessment	Level 2 – Rapid Assessment	Level 3 – Intensive Assessment
General description: Imagery based assessment of landscapes	General description: Rapid site integrity assessment	General description: Detailed site integrity assessment
Evaluates: Integrity of both on and off-site conditions around individual sites/occurrences using indicators within occurrences that are visible with remote sensing data, and Indicators in the surrounding landscape / watershed	**Evaluates:** Integrity of individual areas/occurrences using relatively simple field indicators • Very rapid assessment (narrative) • Rapid assessment (standard metrics) • Hybrid assessments (rapid + vegetation plot)	**Evaluates:** Integrity of individual areas/occurrences using relatively detailed quantitative field indicators. Choice of metrics for detailed assessment may differ from that for monitoring.
Based on: • GIS and remote sensing data • Layers typically include: • Land cover, land use, other ecological types • Stressor metrics (e.g., roads, land use)	Based on: • On-site condition metrics (e.g., vegetation, hydrology, soils,) • Stressor metrics (e.g., ditching, road crossings, and pollutant inputs)	Based on: • On-site condition metrics (e.g., vegetation, hydrology, soils) • Indicators that have been calibrated to measure responses of the ecological system to disturbances (e.g., indices of biotic or ecological integrity)
Potential uses: • Identifies priority sites • Identifies status and trends of acreages across the landscape • Identifies condition of ecological types across the landscape • Informs targeted restoration and monitoring	Potential uses: • Relatively inexpensive field observations across many sites • Informs monitoring for implementation of restoration, mitigation or management projects • Landscape / watershed planning • General conservation and management planning	Potential uses: • Detailed field observations, with repeatable measurements, and statistical sampling design • Identifies status and trends of specific occurrences or indicators • Informs monitoring for restoration, mitigation, and management projects

Level 1 Remote Assessments rely almost entirely on Geographic Information Systems (GIS) and remote sensing data to obtain information about landscape condition and stressors in and around an occurrence. They can also help assess the distribution and abundance of ecological types in the landscape or watershed. **Level 2 Rapid Assessments** use relatively simple field metrics for collecting data on specific occurrences, and will often require considerable professional judgment. Our approach emphasizes a condition-based rapid assessment, supplemented by information on stressors that may be affecting condition. **Level 3 Intensive Assessments** require more rigorous, field-based methods that provide higher-resolution information on the wetland occurring within an assessment area, often employing quantitative plot-based assessment procedures coupled with a sampling design. Calculations

of calibrated indices, such as a Vegetation Index of Biological Integrity (VIBI) may also be used. This 3-level approach to assessments, summarized in Table 3, allows the flexibility for developing data on many occurrences that cannot readily be visited or intensively studied as well as those for which detailed information is desirable. When coupled with standardized procedures for defining occurrences across the landscape (NatureServe generic EO specs), it encourages a widespread application of ecological integrity assessments (assigning EORANKs) based on a reasonable and cost-effective approach for the programmatic or project needs.

The 3-level approach is intended to provide increasing accuracy of ecological integrity assessment, recognizing that not all conservation and management decisions need equal levels of accuracy. At the same time, the 3-level approach allows users to choose their assessment based in part on the level of classification (and thereby the specificity of the conceptual model).

To ensure that the 3-level approach is consistent in how ecological integrity is assessed among levels, a standard framework or conceptual model, such as the EIA Model introduced above, should be used for choosing metrics. Using this model, a similar set of metrics would be chosen across the 3 levels, organized by the standard set of ecological attributes and factors, such as landscape context, size, and condition.

Calibrating the Three-Level Approach

Ideally, information at the three levels of assessment provides relatively consistent information about ecological integrity, with improved interpretations as the level of intensity goes up. To achieve this, the various levels need to be calibrated against each other. For example, sites where a Level 2 or Level 3 IEI or VIBI has been determined can be used to calibrate the Level 1 remote-sensing based index of integrity (Mack 2006, Mita et al. 2007, Fennessy et al. 2007).

Level 1 Assessment (Remote-Sensing Metrics)

Overview

Level 1 Assessments are based primarily on metrics derived from remote sensing imagery. The goal is to develop metrics that assess the landscape context and the on-site conditions of an ecosystem. Satellite imagery and aerial photos are the most common sources of information for these assessments. Typically it is the stressors to the ecological integrity of ecosystems that are most observable with these sources of information, so condition is evaluated through the lens of stressors. Level 1 assessments are widely used as part of regional assessments because of their ability to characterize large landscape areas.

There are growing sets of information on various kinds of stressors that impact ecosystems. Danz et al. (2007) noted that "Integrated, quantitative expressions of anthropogenic stress over large geographic regions can be valuable tools in environmental research and management." When they take the form of a map, or spatial model, these tools initially characterize ecological conditions on the ground; from highly disturbed to apparently unaltered conditions. They can be particularly helpful for screening candidate reference sites; i.e., a set of sites where anthropogenic stressors range from low to high. Ecological condition of reference sites are further characterized to determine how ecological attributes are responding to apparent stressors. This knowledge may then apply in other similar sites.

Anthropogenic stressors come in many forms, from regional patterns of acid deposition or climate-induced ecosystem change, to local-scale patterns in agricultural drainage ditches and tiles, point-source pollution, land-conversion, and transportation corridors, among others. To be effective, a landscape condition model needs to incorporate multiple stressors, their varying individual intensities, the combined and cumulative effect of those stressors, and if possible, some measure of distance away from each stressor where negative effects remain likely. Since our knowledge of natural ecosystems is varied and often limited, a primary challenge is to identify those stressors that likely have the most degrading effects on ecosystems or species of interest. A second challenge is to acquire mapped information that realistically portrays those stressors. In addition, there are tradeoffs in costs, complexity, the often varying spatial resolutions in available maps, and the variable ways stressors operate across diverse land and waterscapes. Typically, expert knowledge forms the basis of stressor selection, and relative weighting. Once models are developed, they may be calibrated with field measurements. Developing empirical relationships between stress variables and ecological response variables is a key to providing insights into how human activities impact ecological condition (Danz et al. 2007).

Two related approaches may be taken to developing Level 1 metrics. First, emphasis can be placed on a comprehensive evaluation of the entire landscape, based on mapped information of stressors. The method compiles and integrates multiple layers of information into an overall synthetic index of landscape and site stressors (as e.g. described by Danz et al. 2007, also Mack 2006) into an overall. The overall index can also be decomposed into individual stressors or sets of stressors, to determine which may be most important. This is a "stressor-based approach," and it assesses ecological integrity of occurrences at specific sites somewhat indirectly.

Second, emphasis can be placed on a method where sites are evaluated using remotes sensing metrics that estimate ecological integrity more directly. The overall index brings together a series of metrics from site, buffer and surrounding landscape that characterize ecological integrity.

For example, NatureServe has developed a Landscape Condition Model (LCM, Comer and Hak 2009), similar to the Landscape Development Index used by Mack (2006). It is a regional GIS model of landscape condition, originally established as a 30m grid of unique values. The algorithm integrates various land use GIS layers (roads, land cover, water diversions, groundwater wells, dams, mines, etc.) (Table 1). These layers are the basis for various stressor-based metrics. The metrics are weighted according to their perceived impact on ecological integrity, into a distance-based, decay function to determine what effect these stressors have on landscape integrity. The result is that each grid-cell (30 m) is assigned a "score". The product is a watershed map depicting areas according to their potential "integrity." The index is segmented into three or four rank classes, from Excellent (minimally disturbed) (A) to Poor (degraded) (D), in accordance with Table A4.

TABLE A4. Example of Level 1 metrics for assessing ecological integrity of an ecosystem, based primarily on "stressor metrics."

Theme
Transportation [1]
Primary Highways with limited access
Primary Highways without limited access
Secondary and connecting roads
Local, neighborhood and connecting roads
Urban and Industrial Development[2]
High Density Developed
Medium Density Development
Low Density Development
Managed & Modified Land Cover[2]
Cultivated Agriculture
Pasture & Hay
Managed Tree Plantations
Introduced Upland Herbaceous
Introduced Wetland Vegetation
Introduced Tree & Shrub
Recently Logged
Native Vegetation with Introduced Species
Ruderal Forest & Upland

[1] A common source of data for these stressors in the U.S. are the ESRI® Data & Maps: StreetMap™ Series issue: 2006 United States, 1:100,000 and National Land Cover Data/ LANDFIRE Existing Vegetation. 2001–2003 United States. 30m pixel/ 1:100,000

NatureServe has also developed a number of Level 1 EIA methods, which uses some of the same information available in the Landscape Condition Model, but selecting measures that are more directly relevant as indicators of integrity, such as buffer extent or connectivity. One version, shown in Table A5, uses fairly simple measures from remote sensing imagery to develop relatively simple level 1 metrics (what can be called "tier 1" metrics). For example, connectivity is measured simply as the proportion of natural land cover in the landscape area around an occurrence. This version of Level 1 EIA is general and may be applicable to all terrestrial ecosystems.

TABLE A5. Example of a general multi-metric approach for a Level 1 Ecological Integrity Assessment.

RANK FACTOR **Metrics** ***Submetrics***	**Weight**
LANDSCAPE CONTEXT	
Connectivity	0.5
Connectivity: % Natural Land Cover in 100 ha area	2
Connectivity: % Natural Land Cover in 1000 ha area	1
Surrounding Land Use Index	0.5
Surrounding Land Use: Score for 100 ha area	2
Surrounding Land Use: Score for 1000 ha area	1
Buffer Index	1
Percent Buffer	1
Average Buffer Width	1
SIZE	
Size	
CONDITION	
On-Site Land Use Index	1

More sophisticated Level 1 metrics can be developed, where the imagery is interpreted in greater detail (we can refer to these as tier 2 metrics) (Table A6). For example, connectivity can be assessed using "Circuitscape" which uses circuit theory to predict connectivity in heterogeneous landscapes for individual movement, gene flow, and conservation planning (see www.circuitscape.org/Circuitscape/Welcome.html). Landscapes are represented as conductive surfaces, with low resistances assigned to habitats that are most permeable to ecological processes such as species movement or best promote gene flow, and high resistances assigned to poor dispersal habitat or to movement barriers (McRae et al. 2008). Similar other metrics, such as Fire Regime, are estimated from information on age class distributions modeled on the landscape. Because these models draw on more detailed information, the metrics can be tailored to more specific sets of ecosystems. For example, the fire regime class metric may have different ratings for systems with broadly different fire regimes. (See full details in Appendix 1.)

TABLE A6. Example of a specific multi-metric approach for a Level 1 Ecological Integrity Assessment for Great Basin Dryland Ecosystems. A full example is provided in Appendix A.5.1.

Rank Factor **Key Ecological Attribute and Specific Indicator**
LANDSCAPE CONTEXT
Landscape Connectivity: CircuitScape Index
Landscape Condition: Landscape Condition Model Index
SIZE
Change in Extent: Relative Area Lost to Land Conversion
CONDITION
Fire Regime: Departure from expected distribution of age classes using SCLASS
Native Species Composition: Invasive Annual Cover

The Role of Level 1 Assessments for Field-based Surveys

The Level 1 integrity ranks are often used as a means of prioritizing sites for field visits, where Level 2 or Level 3 assessments will be completed (e.g., see Fennessy et al. 2007), and ranks based on those assessments would supersede these ranks. Thus level 1assessments can be informative about the overall range in conditions across a population of wetlands in a landscape or region. They can serve as a helpful screening method for identifying the most likely conditions on the ground.

Level 1 ratings can also be used as predictors of Level 2 or 3 ratings at individual sites. Tests completed to date, however, show that Level 1 methods do not accurately predict individual site ratings, particularly on-site conditions (Mack 2006, Fennessy et al. 2007). Our recent tests for wetland site in Michigan and Indiana bear this out (Fig. A4). However, the models are more successful in predicting overall IEI scores, because landscape context and size, as well as on-site condition are relevant to an IEI. It may also be possible to recalibrate the metrics used for Level 1 assessments based on these Level 2 scores.

FIGURE A4. Examples of a correlation between two kinds of Level 1 EIA (remote sensing) models and Level 2 (field-based) assessments of On-site Condition (with scale of A (5.0) to D (D= 1.25). Condition ratings integrate field-based vegetation, soils, and hydrology metrics. The Landscape Condition Model (LCM) uses stressor-based metrics and the Level 1 EIA Method has a combination of stressor and integrity based metrics.

FIGURE A4a) LCM rating within the combined area of wetland and core landscape (1 km buffer around wetland). Kruskal-Wallis F = 32.6, p < 0.001. VL= L< M< H =VH. (from Faber-Langendoen et al. 2011).

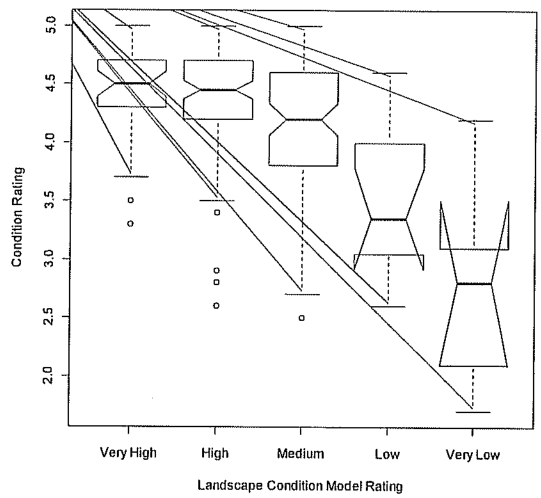
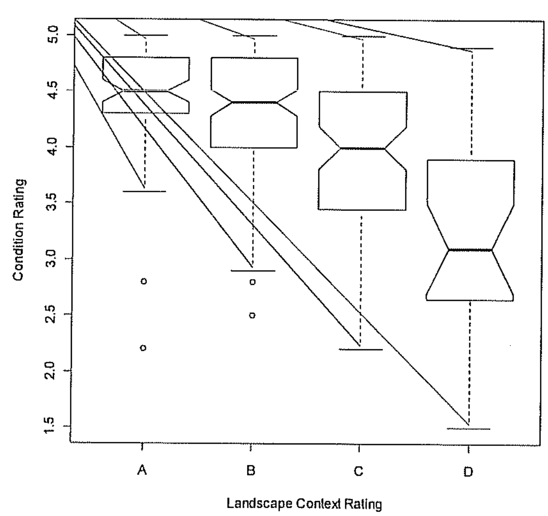


FIGURE A4b) Level 1 EIA Landscape Context ratings, based on three landscape and buffer metrics. Kruskal-Wallis F = 52.0, p < 0.001. A>B>> C>>D (Faber-Langendoen et al. 2011)



Level 2 Assessment (Rapid Field-Based Metrics)

Level 2 Ecological Integrity Assessment Metrics

The intent of ecological integrity–based rapid assessment methods (RAMs) is to evaluate the complex ecological condition of a selected ecosystem using a specific set of observable field indicators, and to express the relative integrity of a particular occurrence in a manner that informs decision-making, whether for restoration, mitigation, conservation planning, or other ecosystem management goals (Stein et al. 2009). These Level 2 assessments are structured tools combining scientific understanding of ecosystem structure, composition, and processes with best professional judgment in a consistent, systematic, and repeatable manner (Sutula et al. 2006).

Metrics that are chosen should informative about integrity or sustainability of major or key ecological attributes and to associated stressors (this is sometimes described as the metrics showing a "stressor-dose response" to changes in stressor levels). Stressor tests can be conducted to ensure that metrics are informative, by assessing how metrics respond to a gradient of stressors levels (Rocchio 2007, Jacobs et al. 2010, Faber-Langendoen et al. 2011).

Level 2 assessments rely primarily on relatively rapid (~2– 4 hour) field-based site visits, but this may vary, depending on the purposes of the assessment. They provide the opportunity to do direct, ground based surveys of ecosystem occurrences. RAMs have particularly widely available for wetlands because of the need for mitigation and restoration tools, and they are in use by many state wetland programs (Fennessy et al. 2007). Typically three to five metrics are identified for each of the ecological factors, each metric relevant to a key ecological attribute.

Examples of the range of metrics that can be completed for rapid assessments for wetlands are provided in Table A7 and A8. NatureServe has developed a Level 2 Ecological Integrity Assessment method for all wetlands in the U.S., with some metrics having variants for certain ecosystem types (using formations and macrogroups) or hydrogeomorphic types (using HGM classes) (see Table A7) (Faber-Langendoen et al. 2008, Faber-Langendoen 2009). EPA has also developed a rapid assessment tool as part of an ambient monitoring program, called USA RAM (Collins and Fennessy 2010 draft) (Table A8). The protocol is similar to that of the NatureServe EIA method, but metrics more often measure degree of complexity or diversity of ecosystems, which may or may not be reflective of ecological integrity.

TABLE A7. Example of a standard set of indicators based on the conceptual model of Ecological Integrity, for wetlands. Indicators occur at different levels of precision (rank factor, major ecological factor, metric). Not shown are how some of the metrics have variants based on wetland type (bog & fen, marsh, floodplain & swamp forests, mangrove, etc.). See Appendix 2 for details on each metric.

RANK FACTOR	MAJOR ECOLOGICAL FACTOR	METRIC NAME
LANDSCAPE CONTEXT	LANDSCAPE CONTEXT	Connectivity (Core, Supporting)
		Land Use Index (Core, Supporting)
		Barriers to Landward Migration (% of Perimeter Obstructed) (tidal)
	BUFFER	Buffer Index • Average Buffer Width • Percent Buffer • Buffer Condition
SIZE	SIZE	Relative Patch Size (ha)
		Absolute Patch Size (ha)
CONDITION	VEGETATION	Vegetation Structure
		Regeneration (woody)
		Native Plant Species – Cover
		Invasive Exotic Plant Species – Cover
		Vegetation Composition
	HYDROLOGY	Water Source
		Hydroperiod
		Hydrologic Connectivity
	SOIL	Physical Patch Types
		Soil Disturbance (Soil Surface Condition)

The checklists provide additional field information on stressors to the wetland site or occurrence. Details are provided in Faber-Langendoen et al. (2008).

TABLE A8. Draft Scoring Criteria for USA-RAM v10.

ECOLOGICAL FACTOR	METRIC NAME
LANDSCAPE	Wetland Abundance
	Landscape Connectivity
BUFFER	Percent of Assessment Area having Buffer
	Buffer Width
HYDROLOGY	Hydroperiod
	Hydrologic Connectivity
PHYSICAL STRUCTURE	Topographic Complexity
	Patch Type Diversity
VEGETATION	Vertical Complexity
	Plant Community Complexity

Level 2 Stressors Checklists

Stressor checklists can be useful as additional information when evaluating the ecological integrity of an occurrence (Kapos et al. 2002). Typically, they are an aid to further understanding the overall condition of the wetland. In some cases, where stressors appear to be having a negative impact on the site, but the condition metrics do not reflect these impacts, it may lead to changes in the overall index of ecological integrity of a wetland. This should only be done in exceptional circumstances. The need for manual overrides may suggest that the current condition metrics may be insensitive to degradation due to certain stressors, and future adjustments to the metrics used may be needed.

Stressors are listed if they are observed or inferred to occur, but are not included if they are projected to occur in the near term, but do not yet occur. Stressors may be characterized in terms of scope and severity. **Scope** is defined as the proportion of the occurrence of an ecosystem that can reasonably be expected to be affected (that is, subject to one or more stresses) by the threat with continuation of current circumstances and trends. Within the scope (as defined spatially and temporally in assessing the scope of the threat**), severity** is the level of damage to the ecosystem from the threat that can reasonably be expected with continuation of current circumstances and trends by excluding potential new threats). For ecosystems, severity is typically assessed by known or inferred degree of degradation or decline in integrity to one or more key ecological attributes.

Standardized checklists of stressors have been developed for a variety of rapid assessment methods (Collins et al. 2006, Faber-Langendoen et al. 2008, Collins and Fennessy 2010). They can be used to create field-based versions of stressor indices, e.g., the Human Stressor Index of Rocchio (2007) integrates stressor scores for hydrology, soils, and buffer.

Variation on the Level 2 Assessment

It is worth noting several variants of the Level 2 EIA assessment methods that may appeal to different needs. First, there is the "very rapid assessment method," in which, the attributes themselves serve as the general indicators, and field crews complete a structured narrative evaluation of those attributes. For example, field crews may record observations on the buffer, vegetation, soils, and hydrology, and then rate the attribute directly. While not preferred as a general, standard method, it remains a valuable approach for professional ecologists, well-experienced in the range of variation in wetland conditions and degradation, provided the reasons being the ratings are documented. It can also be a form of accuracy assessment for Level 1 assessments. This approach has been widely used by the NatureServe network.

A second variant is to complete a standard level 2 method, but add a few select level 3 indicators because it is important for the goals of the project to better understand some key attributes. It may also be desirable to continue collecting detailed information on certain attributes to validate the level 2 assessment. We may refer to these as the "slower rapid assessment method." A common addition is that of a vegetation plot, or some type of standardized plant species list for an occurrence. These data can provide sufficient composition information for VIBIs or FQIs, or structural characteristics (e.g., old growth or coarse woody debris ratings in forests). As long as the modifications are structured within the overall conceptual model, there should be little difficulty in producing comparable results to other RAMs. This approach has also been widely used by the network.

Validation of Level 2 Methods

Because RAMs are rapid-based metrics and rely in part on best professional judgment, it is important that they be calibrated and validated against independent measures of wetland condition in order to establish their scientific defensibility (Sutula et al. 2006, Fennessy et al. 2007). But RAMS often assess a wider range of ecological attributes of wetland integrity (from buffer to hydrol, vegetation and soils, as the EIA conceptual model calls for, so the challenge is to identify a comprehensive set of intensive metrics that span the same range of attributes. Otherwise, calibrations run the risk of optimizing the RAM for only one or several attributes. For example, many level 3 assessments focus on vegetation or biota (e.g., VIBIs), expecting in part that they integrate the response expected from the rest of the biophysical environment. Although this may be true, it too comes with assumptions. Thus, a truly comparable set of level 3 assessments are needed to calibrate level 2 assessments. That said, level 3 assessments of particular attributes can help validate parts of a level 2 assessment. As Stein et al. (2009) point out, decisions regarding modification of RAM components can be made based on a "weight-of-evidence" approach, that is, to combine information from multiple lines of evidence to reach a conclusion. In fact, such an approach is used to select indicators for RAMs in the first place.

Level 3 Assessments (Intensive Field Metrics)

The intent of intensive methods for evaluating ecological integrity is to develop data that are rigorously collected, often with an explicit sampling design, to provide better opportunities to assess trends in ecological integrity over time. The quantitative aspect of the indicators lends themselves to more rigorous testing of the criteria for metric selection (see KEY ECOLOGICAL ATTRIBUTES AND INDICATORS above). Because of their cost and complexity, level 3 methods are often closely evaluated to ensure that they address key decision-making goals, whether for restoration, mitigation, conservation planning, or other ecosystem management goals. They are often highly structured methods, with detailed protocols that ensure a consistent, systematic, and repeatable method (Sutula et al. 2006). The level of intensity required of level 3 methods typically means that they are used in conjunction with level 1 and 2 methods to increase spatial representation and maintain affordability.

As with other levels, metrics that are chosen should be informative about integrity or sustainability of major or key ecological attributes and to associated stressors. Stressor tests can be conducted by assessing how metrics respond to a gradient of stressors levels (Rocchio 2007, Jacobs et al. 2010, Faber-Langendoen et al. 2011).

Level 3 metrics, more so than level 1 and 2, allow for greater specification by ecosystem type. The detailed measures may allow for greater sensitivities in differences among ecosystems in terms of ecological processes, structure or composition. Examples of the range of metrics that can be completed for intensive assessments of forest ecosystems is provided in Appendix 3.

Some intensive assessments have focused on one major attribute, that of vegetation. As with aquatic IBI methods (Karr and others), the approach has been to develop a Vegetation Index of Biotic Integrity (VIBI). Use of plants for a terrestrial (dryland, wetland) IBI makes sense (Mack 2007): plants, and especially vascular plants are large, obvious, important components of terrestrial ecosystems; their taxonomy is relatively well understood and regional and state-specific taxonomic treatments are available; the flora is large and offers numerous potential attributes for the development of a plant IBI; quantitative vegetation sampling methods are well developed and relatively easy to implement in the field, sampling is cost effective and the data sets acquired from such sampling have multiple uses including IBI development, setting mitigation wetland performance standards and supporting wetland permit program decision-making (Fennessy et al. 2002). An example of the individual metrics and descriptions that we will use to develop the VIBI are shown in Table A9.

TABLE A9. Example of a VIBI for Freshwater Wet Meadows and Marshes (Mack 2007,Appendix A).

Description of metrics used VIBI-EMERGENT

Metric	Type	Description
Carex	Richness	Number of species in the genus *Carex*. Note number Cyperaceae species used as a substitute metric for Lake Erie Coastal Marshes
Dicot	Richness	Number of native dicot (dicotyledon) species
Shrub	Richness	Number of shrub species that are native and wetland (FACW, OBL) species
Hydrophyte	Richness	Number of vascular plant species with Facultative Wet (FACW) or Obligate (OBL) wetland indicator status (Andreas et al., 2004)
A/P ratio	Richness ratio	Ratio of number species with annual life cycles to number of species with perennial life cycles. Biennial and woody species excluded from calculation
FQAI	Weighted richness index	The Floristic Quality Assessment Index score calculated using Eq. (7) and the coefficients of Andreas et al. (2004)
% Sensitive	Dominance ratio	Sum of relative cover of plants in herb and shrub stratums with a coefficient of conservatism (C of C) of 6, 7, 8, 9 and 10 (Andreas et al., 2004)
% Tolerant	Dominance ratio	Sum of relative cover of plants in herb and shrub stratums with a C of C of 0, 1 and 2 (Andreas et al., 2004)
% Invasive graminoids	Dominance ratio	Sum of relative cover of *Typha* spp., *Phalaris arundinacea*, and *Phragmites australis*
Biomass	Primary production	The average grams per square meter of clip plot samples collected at each emergent wetland

Field Methods and Protocols

Any discussion of metrics would be incomplete without at least a note that documentation of the rationale and protocols for metrics is vital for their consistent application. The protocols guide the field methods and data collection. They ensure that users understand how the assessment was done, and in the case of monitoring programs, are able to repeat the measurements. See Oakley et al. (2003) for recommendations on content for metric protocols. Examples of metric protocols are available in Faber-Langendoen et al. (2008), Tierney et al. (2010), and in Appendix 6. A data dictionary of ecological integrity metrics is under development.

Ecological Integrity Scorecards

When reporting on the results of assessments, a single index is often desired because it provides an overall measure of the ecological integrity of the ecosystem. To be effective, it needs to be a) grounded in the overall ecological integrity model so that the ecological information is clearly summarized, b) be readily understood by managers and the public, and c) be helpful for decision-making, such as whether wetlands are meeting water quality standards (Fennessy et al., 2007, Jacobs et al. 2010). Having used multiple indicators to get a clear picture of the status of a key or major attribute, it makes sense to use the weight of the evidence across indicators to determine the status of the ecosystem.

Andreasen et al. (2001) outline six characteristics that a practical index of ecological integrity should have:

- Multi-scaled
- Grounded in natural history
- Relevant and helpful (to the public and decision-makers, not just scientists)
- Flexible
- Measurable
- Comprehensive (for composition, structure and function)

We have previously outlined how our method addresses the six characteristics (Faber-Langendoen et al. (2006). The construction of our ecological integrity assessments has been driven by the need to represent major and key attributes of integrity, for which specific indicators or metrics are identified. Thus one key feature of the scorecard is that the integrity of those attributes be summarized through the information available from the metrics. Plus, all of the levels of assessment that we describe are gathering data at the level of ecological factors, so scorecards will also retain a common set of levels for reporting, a desirable feature of scorecards highlighted by Harwell et al. (1999). This will also make it easy for users to find the specific indicator information on which these summary scores are based.

The IEI reports on the condition of a wetland by estimating the degree to which individual sites have departed from reference standard conditions. The specific indicators and the attributes are scaled based on best understanding of the range of natural and other variability relevant to ecological integrity, and with reference to sites where the highest scores reflect reference standard conditions and the lowest scores representing highly disturbed sites (Stoddard et al. 2006).

It may also be desirable to summarize the level of stressors present on the site, in order to indicate what might be driving the current levels of integrity. In many cases, these data may only provide correlating support for the level of integrity, and further studies, or supporting evidence from other studies, will be needed to demonstrate causation.

The degree to which the scorecard might need to be customized based on several aspects of ecosystems. First, the relative importance of attributes may differ among system types. For example, hydrology typically plays a very strong role in wetlands, but a minor one in drylands; within wetlands, buffer may play a larger role in depressional wetlands than in riverine systems (Jacobs et al. 2010). Or size may be relatively unimportant to wetland types, such as seeps, that only ever occur in small patches. Second, and related to the above, failure of certain attributes may lead to the overall collapse of a system. Thus more weight can be given to poor ratings of those key attributes. Here again, ecosystems classifications play a valuable role in providing guidance on our understanding of the role of ecological attributes among systems, and can ensure standardized evaluations whenever that ecosystem is encountered. Nonetheless, it may be tempting to focus on the individuality of ecosystems at the expense of readily interpretable results; customizing should be done only where strong and range-wide evidence that it improvise the discriminating ability of the index, to ensure that scores for sites can be readily compared across watershed, landscapes and regions.

There are a number of approaches to aggregating metrics, but the most common is the rather simple **non-interaction point-based approach**, where each metric is scored and treated independently. The point-based approach is consistent with that of many IBI scoring methods (e.g. Karr and Chu 1999). The scorecard is structured using the conceptual model (Fig. A3 above). Each metric within an ecological factor or attribute is assigned a weight, based on its perceived importance. Ratings for each metric are presented, along with conversion to a point value for that rating (e.g., A = 5 points, B = 4, C=3, D=2, E =1). The points are multiplied by the weight to get a score for the metric. The scores (weighted points) for all metrics within a major attribute are summed and divided by the sum of the weights to get an attribute score. Each major factor is also weighted (e.g., in wetlands, soils are often weighted less than either hydrology or vegetation). The factor scores can be further aggregated by major rank factors (landscape context, size, and condition). Finally their scores can be weighted and summed to get an overall score, which is converted to an Index of Ecological Integrity. If desired, Vegetation, Soils and Hydrology can be combined separately into a Condition score before producing an overall index rating. An example of a summary scorecard is shown in Table A10. More detailed examples are provided in Appendices 1, 4, and 5.

TABLE A10. Example of an Ecological Integrity Scorecard.

RANK FACTOR	MAJOR ECOLOGICAL FACTOR / *Metric*	Rating
ECOLOGICAL INTEGRITY		**B**
LANDSCAPE CONTEXT		**B**
	LANDSCAPE	B
	Connectivity	*A*
	Land Use Index	*B*
	BUFFER	B
	Buffer Index	*B*
SIZE		**A**
	SIZE	A
	Relative Patch Size (ha)	*B*
	Absolute Patch Size	*A*
CONDITION		**B**
	VEGETATION	B
	Vegetation Structure	*C*
	Regeneration (woody)	*C*
	Native Plants – Cover	*B*
	Invasive Exotic Plants – Cover	*C*
	Increasers – Cover	*B*
	Vegetation Composition	*B*
	HYDROLOGY	C
	Water Source	*C*
	Hydroperiod	*C*
	Hydrologic Connectivity	*B*
	SOIL	B
	Physical Patch Types	*B*
	Soil Disturbance	*B*

A.7. Definition of Ecological Integrity Rating Values (A–D)

Definitions

The ecological integrity scorecard will bring together our understanding of current status of ecological attributes, and include threshold values for both the best conceivable occurrences and those having only fair viability or integrity (NatureServe 2002). To ensure that the final ratings of ecological integrity have consistency wherever they are used, we provide a narrative summary of the different levels of integrity. Thus, rising from the midst of the ratings and rollups should emerge a sense of integrity that meets these definitions (Table A11).

We offer these definitions partly to provide a global perspective on ecological integrity. This means that the best occurrence in a particular jurisdiction or geographic area (*e.g.,* ecoregion) may not be highly ranked or even viable. Information about local prioritization of EOs can be recorded in optional fields or existing comment fields.

The A through D rating presumes that a particular type is still recognizable at some level as "the type," despite varying levels of degradation. At some point, a degraded type will "cross the line" (or be "transformed" in the words of SER 2004) into a separate, typically semi-natural or cultural type. In some state-and-transition models these may be treated as shifts to an "alternative state." As a matter of practicality, the current system has been lost. This requires working with a set of diagnostic classification criteria, based on composition, structure, and habitat.

Recap of Natural Variability and Ecological Integrity

A few final observations are in order regarding the role of natural and historic variability in informing these definitions of ecological integrity. These are provided to highlight both the importance and limits of using the range of natural or historic variability.

An "A" rank need not be comparable to historical conditions. For example, bison in native Great Plains prairies will not conceivably exist again in their historical condition with herds numbering in the millions, but nevertheless a range of prairie occurrences with *e.g.,* managed herds of differing sizes and conditions, might still be reasonably achievable. In other words, it is still necessary to conceive of a range of integrity, although the range is truncated when compared to EORANK specifications that would have been written 150 years ago. (NatureServe 2002)

The "A" rank threshold should not be based solely on historical information because: a) historical status often cannot be achievable; b) use of historical information could drastically truncate the rank scale for current EOs; and c) historical information is often not known (NatureServe 2002). Nonetheless, for ecosystems, the A-ranked threshold should be based on a "minimally disturbed" reference state, whenever possible (Stoddard et al. 2006).

In order to set a threshold that is reasonably and conceivably achievable for "A"-ranked occurrences, it is necessary to consider restorability so that the threshold is not limited to EOs that are extant (NatureServe 2002).

TABLE A11. Definition of Index of Ecological Integrity values (Faber-Langendoen et al. 2009c).

Rank Value	Description
A	Occurrence is believed to be, across the range of a type, among the highest quality examples with respect to key ecological attributes functioning within the bounds of natural disturbance regimes. Characteristics include: the landscape context contains natural habitats that are essentially unfragmented (reflective of intact ecological processes) and with little to no stressors; the size is very large or much larger than the minimum dynamic area ; vegetation structure and composition, soil status, and hydrological function are well within natural ranges of variation, exotics (non-natives) are essentially absent or have negligible negative impact; and, a comprehensive set of key plant and animal indicators are present.
B	Occurrence is not among the highest quality examples, but nevertheless exhibits favorable characteristics with respect to key ecological attributes functioning within the bounds of natural disturbance regimes. Characteristics include: the landscape context contains largely natural habitats that are minimally fragmented with few stressors; the size is large or above the minimum dynamic area, the vegetation structure and composition, soils, and hydrology are functioning within natural ranges of variation; invasives and exotics (non-natives) are present in only minor amounts, or have or minor negative impact; and many key plant and animal indicators are present.
C	Occurrence has a number of unfavorable characteristics with respect to the key ecological attributes, natural disturbance regimes. Characteristics include: the landscape context contains natural habitat that is moderately fragmented, with several stressors; the size is small or below, but near the minimum dynamic area; the vegetation structure and composition, soils, and hydrology are altered somewhat outside their natural range of variation; invasives and exotics (non-natives) may be a sizeable minority of the species abundance, or have moderately negative impacts; and many key plant and animal indicators are absent. Some management is needed to maintain or restore[7] these key ecological attributes.
D	Occurrence has severely altered characteristics (but still meets minimum criteria for the type), with respect to the key ecological attributes. Characteristics include: the landscape context contains little natural habitat and is very fragmented; size is very small or well below the minimum dynamic area; the vegetation structure and composition, soils, and hydrology are severely altered well beyond their natural range of variation; invasives or exotics (non-natives) exert a strong negative impact, and most, if not all, key plant and animal indicators are absent. There may be little long-term conservation value without restoration, and such restoration may be difficult or uncertain.[8]

[7] By ecological restoration, we mean "the process of assisting the recovery of an ecosystem that has been degraded, damaged, or destroyed... Restoration attempts to return an ecosystem to its historic trajectory" (SER 2004). Restoration may be distinct from rehabilitation, reclamation, creation, mitigation, or ecological engineering, unless they have as part of their goal a restoration as defined above (see SER 2004 for details).

[8] D-ranked sites present challenges. For example, with respect to classification, a degraded type may bear little resemblance to examples in better condition. Whether a degraded type has "crossed the line" ("transformed" in the words of SER 2004) into a new semi-natural or cultural type is a matter of classification criteria. Here we include D ranked examples as still identifiable to the type based on sufficient diagnostic criteria present.

A.8. Returning to the Watershed or Landscape Scale

A clear and consistent scorecard at the site level that is readily repeatable for occurrences everywhere leads us back to the role of ecological integrity assessment methods at larger spatial scales. The IEI can be used to report on the status of ecosystems across watersheds or landscapes (Fig. A5). Jacobs et al. (2010) note how individual indicators for wetlands can easily be relayed to the public or environmental managers to communicate the status of certain kinds of wetland types in the Nanticoke River watershed. Similarly, Faber-Langendoen et al. (2011) use a scorecard approach to report on the overall integrity of individual wetlands across an ecoregions (Fig. A6). This information can then be used to direct management efforts in the appropriate areas to improve the condition of types in a watershed.

FIGURE A5. Ecological Integrity Scorecard Roll-up for Central Basin and Range ecoregions in the southwest United States.

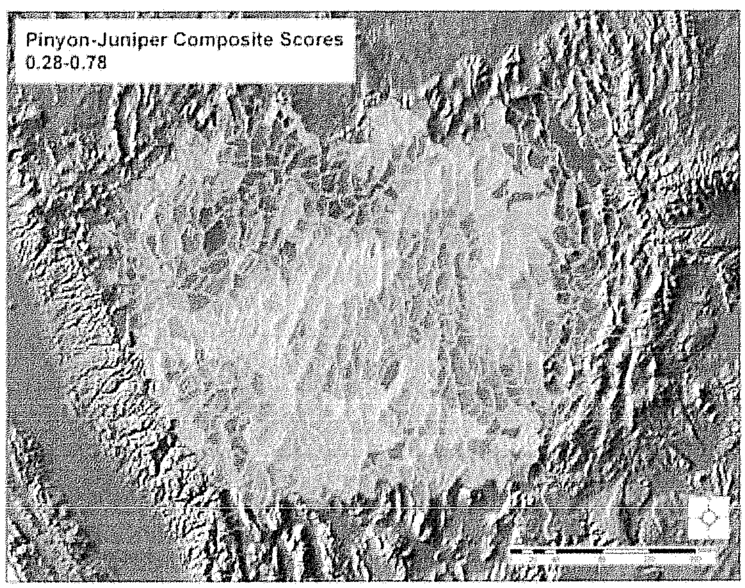
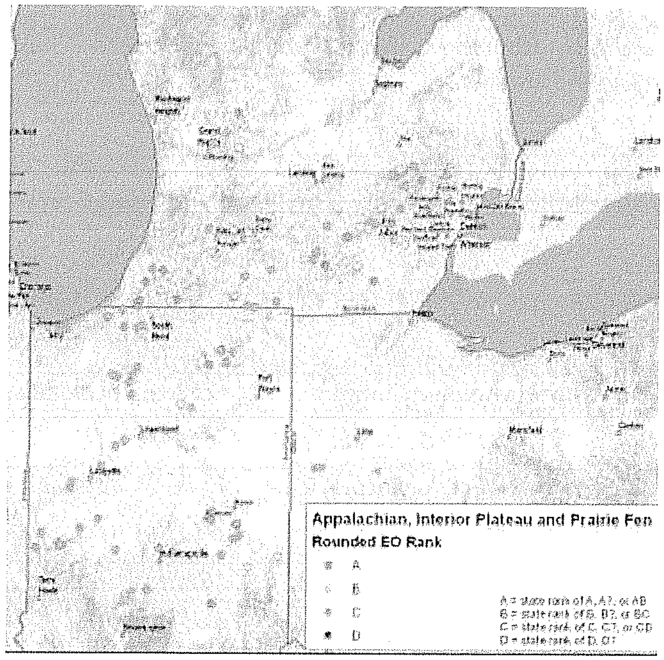


FIGURE A6. Ecological Integrity Scorecard Roll-up for ecoregions of Omernik Ecoregions in northern Indiana and southern Michigan.



A.9. Ecological Integrity, Conservation Status, and Ecosystem Services

Ecological integrity is only one aspect of interest for ecosystem assessments. Two common ones include: 1) conservation (at-risk) status of ecosystem types and biodiversity value, which includes aspects of wetland irreplaceability, and 2) functional values and ecosystem services (Hruby 2001, Fennessy et al. 2004). The first aspect, assessing the conservation status and irreplaceability value of ecosystem types and occurrences, can be part of a risk assessment process, where more irreplaceable systems are preferentially targeted for threat abatement or subject to greater degree of protection, thereby avoiding losses that lead to challenging mitigation or restoration efforts. This assessment can begin by assessing the relative conservation status (or risk of extirpation) of a given type. For example, the Heinz Center (2002) uses the "At-risk wetland plant communities" (based on NatureServe's conservation status assessment approach), as an indicator of overall wetland or aquatic condition.

The second aspect, that of wetland functional or ecosystems services value, has been widely developed as part of the functional assessments completed by the hydrogeomorphic (HGM) approach. Functional assessments and EIAs differ in important respects (Table A12), even as they share many field data methods. Functional assessments categorize wetland types by creating a number of broad wetland classes based on hydrogeomorphic characteristics, then allowing regional applications to specify subclasses. We suggest that the indicators and ecological attributes developed for ecological integrity assessments will often be useful for functional assessments, though additional data may need to be collected. For example, in a functional assessment, a series of measures (e.g., litter + O-horizon thickness + coarse woody debris + snags) are combined with flooding frequency to estimate the degree to which a wetland exports organic carbon, whereas in an ecological integrity assessment, these measures would be combined into abiotic and biotic metrics that assess departures from the ranges of natural variability and characteristics of benchmark sites for these ecosystems.

TABLE A12. Comparison of Ecological Integrity (Condition) and Functional Wetland Assessments. See Fennessy et al. (2007) for further comparisons between the two approaches.

	Ecological Integrity / Condition Assessment	**Functional Assessment**
Purpose	Estimate current ecological integrity	Estimate ecological functions (HGM)
"Currency"	Condition of major and key ecological attributes	Level of functions and ecological services
Approach	"Holistic" ecological integrity	"Compartmental;" each function assessed individually
Method	Combines indicators into conceptual model of ecological factors and key and key ecological attributes	Combines indicators into conceptual model of ecological functions and values
Application	Mitigation, monitoring, state water quality standards, and NatureServe network	Mitigation and monitoring

HGM and other ecosystem services methods may also differ from ecological integrity assessments in that they evaluate the level or capacity of wetland functions, rather than the status of key ecological attributes. They may be concerned with the level or capacity of each function regardless of how or whether it relates to ecological integrity. Thus, a wetland with excellent integrity will perform all of its functions at levels expected for its wetland class or type, whether or not these are optimal levels with respect to desired functions or ecosystem services. Many HGM functional assessments collect very similar data to an ecological integrity assessment; what differs is that the functional assessments may take these data and develop logical operators to infer function.

To enhance the collection of both ecological integrity and ecosystem services, measures can be added to the field or remote sensing methods. For example, National Wetland Inventory maps include a wetland type classification (Cowardin et al. 1985), and more recently include an NWI+ application, which describes the hydrologic functions of the watershed and site scales (USFWS 2010).

A.10. Adapting the Assessment Over Time

We conclude by noting that our efforts to assess ecological integrity are only approximations of our current understanding of the system. Ecosystems are far too complex to be fully represented by a suite of metrics and attributes. Moreover, our metrics, indices and scorecards must be flexible enough to allow change over time as our knowledge grows. What is important is that we present as clearly as we can how we are conducting our assessments, so that we foster communication and understanding among people with different backgrounds, goals, and points of view.

NatureServe will upgrade its databases to manage and store the ecological assessments, including the component metrics, and will encourage new version of metrics to be developed and substituted for old ones as they become available. Programs and partners will be encouraged to test and refine these metrics, keeping in mind the overall definitions and purposes of ecological integrity assessments.

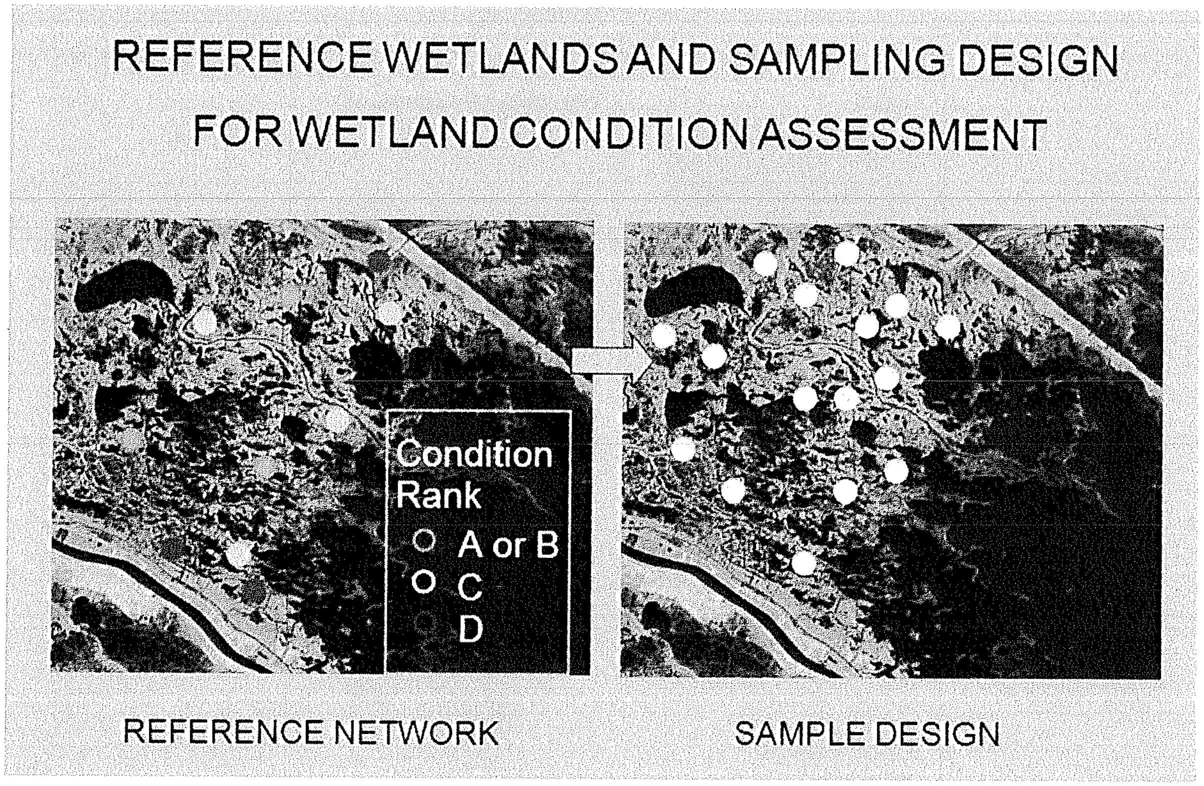
Section B: Identifying a Reference Gradient of Wetland Type and Condition

B.1. Introduction

In this section, we develop a sampling design that will help us identify a candidate set of wetlands that span the ecological gradients in northern Indiana and southern Michigan (all of Omernik level-3 ecoregions 55, 56, and 57, and the Indiana part of 54). We use a sampling design based on previously classified sites, remote sensing metrics and on-site evaluations to predict a reference gradient of conditions (from minimally disturbed to degraded) for each wetland type. This design step is critical, because we needed a sampling design that allowed us to test the relative sensitivity of our Ecological Integrity Assessment (EIA) method to changes in ecological integrity across the full range of wetland types, conditions and stressors (as detailed in Section C).

A follow-up application presented itself through the development of the sampling design. That is, if we successfully create a sampling design that we hypothesize will span a range of conditions, and we independently verify those conditions, then we can use that design to predict the reference gradient for future studies. Predicting a reference gradient will be of great value for monitoring and assessment programs (Fig. B1). This is because, in selecting and establishing metrics for assessing condition or ecological integrity, an assumption is made that some type of reference condition can be defined; that is, it is possible to describe a series of states of wetland integrity, from minimally disturbed to degraded (Stoddard et al. 2006). But there are challenges to implementing the reference condition approach. First, one needs some means of establishing the gradient of reference sites. Second, one needs to be able to identify sites where the range of reference conditions are found; this can be problematic in regions where there is a long and extensive history of human uses that have altered the landscape. Third, one needs to be able to sample these sites in a timely and cost-effective manner to gather the data on reference condition. For that reason, a feasible approach is needed to establish reference conditions and identify candidate reference sites (Faber-Langendoen et al. 2009b).

FIGURE B1. A reference network of wetland sites where ecological integrity has been verified (left panel), which serves to guide a sampling design to assess wetland conditions in a watershed or region (right panel).



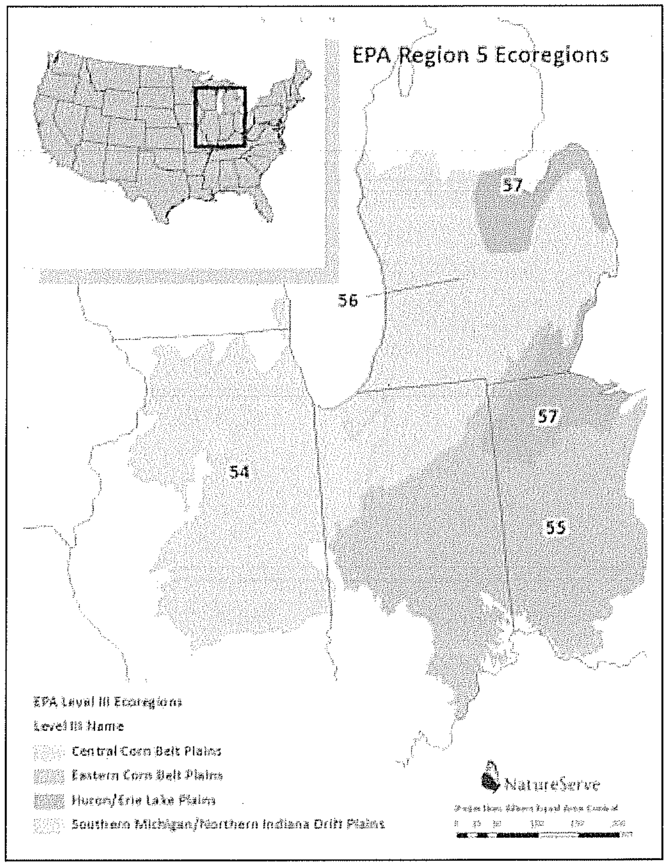
Our purpose in this section is to summarize our proposed sampling design for establishing a reference gradient. Our level of inference for this design and the subsequent testing of our ecological integrity methods will be the range of conditions across a set of occurrences of all major wetland types in southern Michigan and northern Indiana. However, if successful, we further believe that the range of types and conditions will be applicable across the temperate regions of the U.S. and elsewhere.

B.2. Sampling Design Methodology

Project Area

Our sampling area in southern Michigan and northern Indiana is based on the Omernik ecoregions, which are also being used in the EPA National Wetland Condition Assessment (Fig. B2). We conducted our sampling within a relatively similar set of adjacent ecoregions 55, 56, and 57, with a few samples from ecoregions 54 in northwest Indiana.

FIGURE B2. Ecoregional framework that guides the project study area within EPA Region 5. Shown are the Omernik ecoregions (Omernik 1987, updated by USEPA 2007) that were sampled within Lower Michigan and northern Indiana.



This project was coordinated by NatureServe, with field crews comprised of Michigan and Indiana program staff. EPA's National Wetland Condition Assessment (NWCA) Team provided guidance and support.

Classification

The success of developing and assessing wetland ecological integrity depends on understanding the structure, composition, and processes that govern the wide variety of ecosystem types. Ecological classifications can be helpful tools in categorizing this variety. They help ecologists to better cope with natural variability within and among types so that differences between occurrences with good integrity and poor integrity can be more clearly recognized. Classifications are also important in establishing "ecological equivalency," i.e., ensuring that degraded examples of a type are compared with minimally disturbed examples of the same type. We use a variety of classifications in order to effectively address biotic and abiotic aspects of wetlands, at different cales, but our primary focus is on the USNVC macrogroup and the more finely scaled State Natural Community Type. We link these types to NatureServe's Ecological Systems and to Hydrogeomorphic Types.

USNVC Macrogroup

In the United States, the **U.S. National Vegetation Classification** (NVC) is supported by the Federal Geographic Data Committee (FGDC), NatureServe, and the Ecological Society of America, with other partners (FGDC 2008, Faber-Langendoen et al. 2009a, Jennings et al. 2009). The NVC was developed to classify both wetlands and uplands and identify types based on vegetation composition and structure and associated ecological factors. At the highest level of the classification, Formation Class, there are 8 broad classes, each with seven nested hierarchical levels, which permit resolution of types from broad-scale formations to fine-scale associations (Table B1). We use the Macrogroup level for our assessments, of which there are seven in our region (Table B2).

TABLE B1. The following table illustrates the eight levels of the USNVC hierarchy for a Midwest prairie fen. Also shown is an example of how NVC is a complementary classification with Ecological Systems (mid-scale) and Natural Community (finer-scale) types.

USNVC Hierarchy	**Pilot NVCTypes**
Upper Levels	
Formation Class	Low Shrubland & Grassland
Formation Subclass	Temperate & Boreal Shrubland & Grassland
Formation	Temperate & Boreal Bog & Fen
Mid-Levels	
Division	North American Bog & Fen
Macrogroup	Appalachian, Interior Plateau and Prairie Fen
Group	North-Central Appalachian & Interior Seepage Fen
Lower Levels	North-Central Interior Shrub-Graminoid Alkaline Fen System
Alliance	Shrubby-cinquefoil / Fine-leaved Sedges Fen Alliance Prairie Fen (Michigan); Fen (Indiana)
Association	Shrubby-cinquefoil / Sterile Sedge - Big Bluestem - Indian-plantain Fen Vegetation

TABLE B2. List of formations and macrogroups found in the study area, and number of study sites found for each macrogroup. Colloquial names for macrogroups are provided in square brackets.

NVC FORMATION Macrogroup
SWAMP & FLOODED FOREST
Northern & Central Floodplain Forest (M029) **[Floodplain Forest]**
Northern & Central Swamp Forest (M030) **[Swamp Forest]**
FRESHWATER SHRUBLAND, WET MEADOW & MARSH
Eastern North American Wet Shrub, Meadow & Marsh (M069) **[Wet Shrub, Meadow & Marsh]**
Great Plains Wet Meadow, Wet Prairie & Marsh (M071) **[Wet Prairie]**
Atlantic & Gulf Coastal Plain Pondshore & Wet Meadow (M067) **[Coastal Plain Pondshore]**
BOG & FEN
Appalachian, Interior Plateau & Prairie Fen (M061) **[Rich Fen]**
North American Boreal Bog & Fen (M062) **[Bog & Poor Fen]**

State Natural Community Type

Many states throughout the eastern United States have developed natural community classifications with a focus on inventory and conservation applications, including Indiana (Jacquart et al. 2002, see also *www.in.gov/dnr/nature preserve/4743.htm*) and Michigan (Kost et al. 2007). They rely on a suite of state-level ecological characteristics, such as vegetation physiognomy, species composition, soil moisture, substrate, soil reaction, or topographic position, to identify the type. State natural community types are very comparable to alliances and associations in the USNVC (Table B1). We include the state classifications by macrogroup, because states report their wetland information based on these types, and they help summarize the range of variation of wetland types within macrogroups.

Other Classifications

Ecological Systems

A second, related classification approach to that of the USNVC is the **Ecological Systems** classification[9] (Comer et al. 2003). It can be used in conjunction with the USNVC, roughly corresponding to the "group" level, and below the macrogroup level (Table B1, Table B3).

[9] Ecological Systems in the U.S. are a component of the **International Terrestrial Ecological Systems Classification** (Comer et al. 2003).

Ecological Systems provide a spatial-ecologic perspective on the relation of associations and alliances (fine-scale plant community types), integrating vegetation with natural dynamics, soils, hydrology, landscape setting, and other ecological processes. The Ecological Systems classification facilitates mapping at meso-scales (1:24,000–1:100,000). Comprehensive Ecological Systems maps are available across the country (Comer and Schulz 2007, www.landscope.org). We use Ecological Systems maps to help identify where various macrogroups may be found across the landscape, and to characterize the landscape surrounding wetlands.

TABLE B3. Macrogroups and Ecological Systems in the study area.

Macrogroup	Ecological System
Northern & Central Floodplain Forest	North-Central Interior Floodplain (Laurentian-Acadian Floodplain Forest)
Northern & Central Swamp Forest	Laurentian-Acadian Alkaline Conifer-Hardwood Swamp North-Central Interior and Appalachian Rich Swamp North-Central Interior Wet Flatwoods
Atlantic and Gulf Coastal Plain Pondshore and Wet Meadow	Northern Atlantic Coastal Plain Pond
Eastern North America Wet Shrub, Meadow & Marsh	Laurentian-Acadian Freshwater Marsh Laurentian-Acadian Wet Meadow-Shrub Swamp Great Lakes Wet-Mesic Lakeplain Prairie Northern Great Lakes Interdunal Wetland Great Lakes Freshwater Estuary and Delta Northern Great Lakes Coastal Marsh
Appalachian, Interior Plateau & Prairie Fen	North-Central Interior Shrub-Graminoid Alkaline Fen
North American Boreal Bog & Fen	Boreal-Laurentian Bog Laurentian-Acadian Alkaline Fen Boreal-Laurentian-Acadian Acidic Basin Fen

Hydrogeomorphic Classification

The hydrogeomorphic (HGM) classification developed by Brinson (1993) was developed in order to assist the U.S. Army Corp of Engineers with the evaluation of wetland impacts. HGM identifies groups of wetlands that function similarly, based on three fundamental factors: geomorphic setting, water source, and hydrodynamics (Smith et al. 1995). HGM classifications are widely used by wetland scientists. For each wetland occurrence that we visited, we assigned the HGM class.

Wetland Occurrence Data – Natural Heritage and Other Datasets

We first needed to establish our overall population of wetland sites. We assessed the following sources of information:

- State program site information, in which the wetland type (determined using state type, NatureServe ecological system type, and NVC macrogroup type), location (typically a point and a polygon), and condition evaluation, are available
- All Michigan Rapid Assessment Method (MiRAM) sites where sufficient information was available to determine macrogroup, and approximate condition, and functional rating
- The Ecological Systems map for the region, showing all major wetland sites, classified by the NatureServe Ecological Systems types

Natural Heritage Datasets

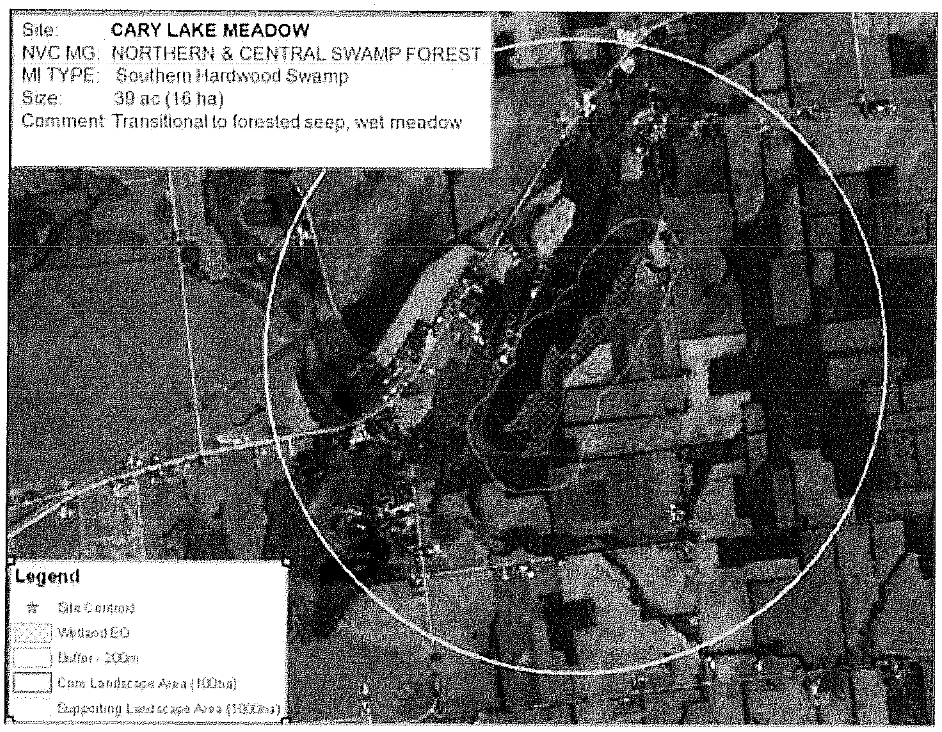
The natural heritage programs in Michigan and Indiana have for many years been identifying high-quality (minimally disturbed) occurrences for all wetland types in the state (using state natural community classifications), and for rarer types, a fuller range of sites. No other comparable datasets are available in these states; in addition, many other states have programs with comparable data. Network programs have typically evaluated the condition of occurrences using best professional judgment, with a minimal amount of quantitative information, assigning each occurrence a scorecard grade, or EORANK: A (Excellent), B (Good), C (Fair), and D (Poor). Programs may differ in how these grades or ranks are assigned, but typically the primary focus was the on-site condition of the wetland.

The natural heritage program databases contain the following core information on each occurrence:

- Site Name
- State Natural Community / NVC Association name
- Element Occurrence ID (unique database code)
- EO Rank
- Geo-coordinates
- Directions to Site
- Polygon delimiting the extent of the wetland type occurrence
- Size of polygon
- Ownership
- Date of visit(s)
- Occurrence Comments (general description of occurrence, including dominant species)

An example of the spatial information and some of the core fields of information is shown in Figure B3.

FIGURE B3. A typical example of the kind of tabular and geospatial information available from network programs. The hatched blue area shows a polygon of a Southern Hardwood Swamp (state type): Northern & Central Swamp Forest (NVC macrogroup), covering 39 ac (16 ha). The Element Occurrence ID is 3473, and the project ID is 74. Assigned rank was "B." Circular areas are landscape areas evaluated as part of the landscape context, including buffer (blue line), core landscape (red line) and supporting landscape (yellow).



To compile this data into our framework, we crosswalked each state type to the USNVC macrogroup level (Table B1). State types are relatively fine-grained relative to macrogroups, so state types typically nest cleanly with the NVC macrogroup; thus our confidence in classification of sites at the macrogroup level is very high. We could then bring in the EORANK data, and compile a list of wetland occurrences by macrogroup and EORANK across the several EPA ecoregions of interest in southern Michigan and northern Indiana. More than 500 wetland occurrences were found across the region.

MiRAM Sites

MiRAM assesses overall wetland function and condition at a wetland site, across multiple wetland types (MDEQ 2008). However, the MiRAM score for the entire wetland is typically a good reflection of the condition of an individual type within the bigger wetland (T. Losee pers. comm. 2009). Data on the location of MiRAM sites was available as point coordinates (hardcopy maps showing wetland polygons were also available).

MiRAM sites were assigned to a macrogroup based on descriptive text recorded for each site; in our initial data analysis, we were unable to assign a macrogroup for 48 of 68 MiRAM sites due to insufficient descriptive information. Only those MiRAM sites with assigned macrogroups were considered for the site selection.

To determine how best to use the MiRAM information on wetland condition, we compared Michigan EO rankings to the MiRAM score for 18 MiRAM sites that are co-located with EOs. MiRAM scores include a Floristic Quality Index (FQI). The FQI proved to be the best match to EORANK. We found that the range of FQI values provided in Table B4 were the best means to split those values into condition classes comparable to the EORANK; the match to the EO rank was generally close, providing adequate means to assign MiRAM sites to both macrogroup and condition ranks.

TABLE B4: Condition Ranges Used to Stratify Project Data.

Original Condition Ranking			Project Condition Ranking
Michigan EORANK	Indiana EORANK	MiRAM FQI (all) Score	
A, AB, A?	A, AB	FQI ≥ 45	A
B, BC, B?	B, BC, AC	35 ≤ FQI < 45	B
C, C?	C, BD	25 ≤ FQI < 35	C
CD, D	CD, D	FQI < 25	D
X, H, ?	X, H, ?	NA	Not Used

Ecological Systems Maps

NatureServe has comprehensive Ecological Systems maps across the project area (and country) (Comer and Schulz 2007, *www.landscope.org*). However, wetland types were typically mapped very generally, above the macrogroup level, so its accuracy is not strong enough at fine-grained scales for our purposes. These maps are better used at local catchment to watershed scales to identify wetlands and types. Given this, we did not rely on the Systems maps for this project.

Sampling Design for Reference Gradient

Classification and Condition Strata

We used classification stratum and condition stratum for site selection and analysis. Our *classification stratum* is the NVC wetland types at the macrogroup level. Our crosswalk from state natural community types to the macrogroup level was very clean, so for most sites we have a high confidence in the classification stratum.

Our second stratum is the *condition stratum,* which predicts the potential integrity or condition of an occurrence. Here our confidence is not as strong, as EORANKS are sometimes outdated, and are based on the best professional judgment of field ecologists who visited the site, which may vary among ecologists and states. Nonetheless having pre-existing field observations is very valuable. To increase the standardization of the grade, we combined the on-site EORANK with a remote-sensing based landscape-context evaluation. Three primary metrics were used: naturalness of surrounding landscape, land uses within the landscape, and the extent and condition of the buffer immediately surrounding the wetland (Table B5). Together they contribute to an overall Landscape Context Rating (LC) rating. Details of the metrics are provided in Appendix 3.

TABLE B5. Landscape Context metrics used to develop a rating around each natural heritage occurrence. See also Figure B3.

Metric **Submetric**	**Weight**
LANDSCAPE CONTEXT	
Connectivity	0.5
Connectivity: % Natural Land Cover in core 100 ha area	2
Connectivity: % Natural Land Cover in supporting 1000 ha area	1
Surrounding Land Use Index	0.5
Surrounding Land Use: Score for core 100 ha area	2
Surrounding Land Use: Score for supporting 1000 ha area	1
Buffer Index	1
Percent Assessment Area with Buffer	1
Average Buffer Width	1

The overall score of these landscape metrics were then combined with the on-site condition evaluations to assign each wetland occurrence to a condition stratum i.e., landscape context rating + on-site EORANK = condition stratum rating). We first combined the ratings for landscape context and EORANK as follows:

Condition Stratum Design 1

A = both EORANK and LC = A

B = EORANK and LC or LC and EORANK = A andr B, A and C, or B or B

C = A and D, B and C, B and D, or C and C

D = C and D, D and D

We also later tested a less stringent version of this design (Condition Stratum 2) to see if future applications might benefit from this version:

Condition Stratum Design 2

A = both EORANK and LC1 = A, or A or B

B = EORANK and LC or LC and EORANK = A and C, B and B, or B and C

C = A and D, B and D, or C and C, or C and D

D = D and D

Site Selection and Sample Size

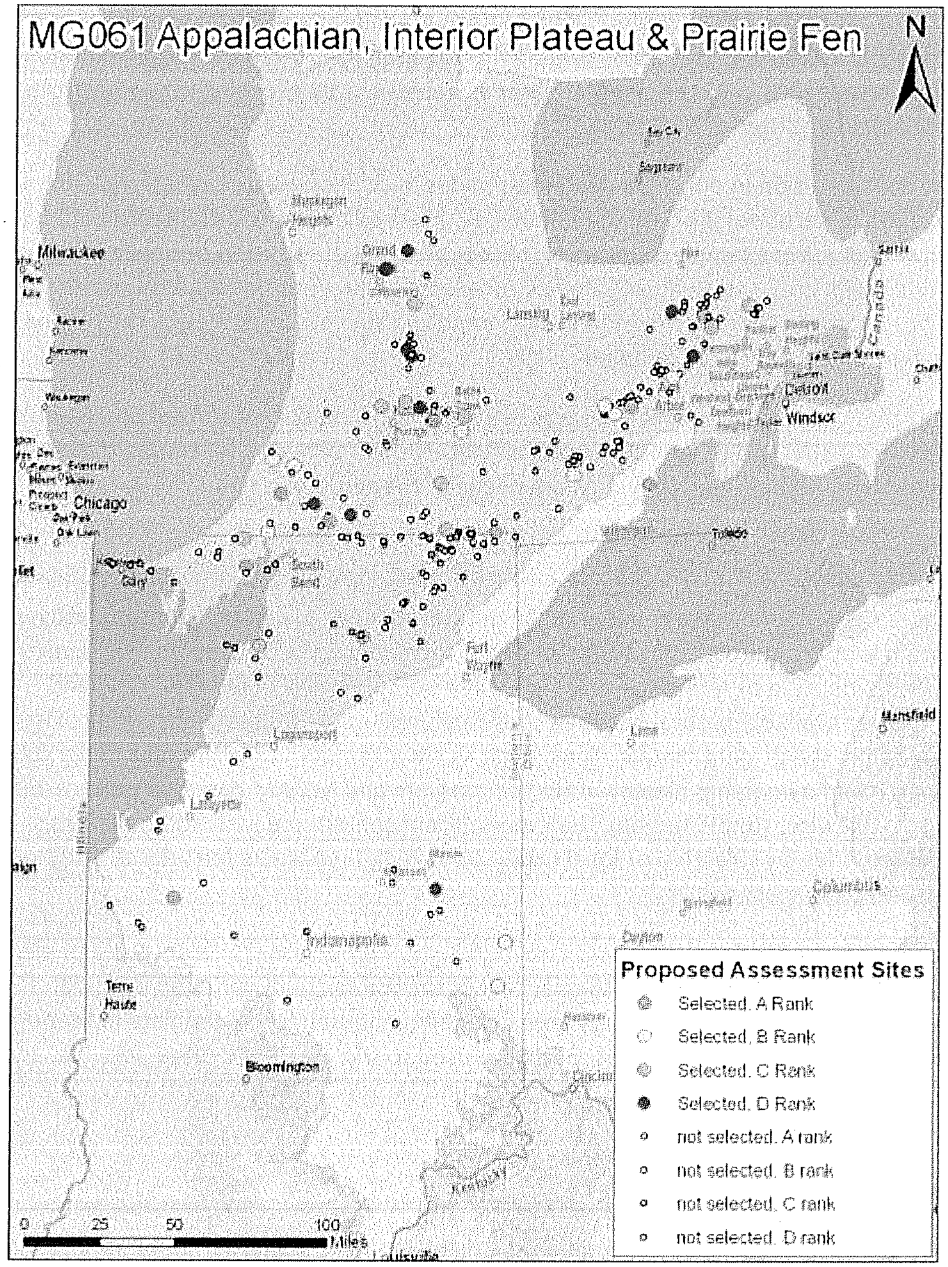
As described above, reference sites were chosen primarily from state programs' element occurrence records for wetland communities, supplemented as necessary with data from MiRAM. We assigned a macrogroup and a condition rating based on the methods described for those two strata above. Table B6 summarizes our achieved number of sampling sites by classification and condition strata, which varies considerably from our initial target of 10 sites per cell. Although our goal was to achieve a balanced a design to ensure that the testing of the EIA method spanned the full range of wetland types and condition, neither the statistical analyses nor our overall interpretation depend on having exactly 10 replicates. Thus, rather than engage in an effort to balance the cells, we emphasized attaining 5 or more sites per cell. In addition, some types (e.g., Bog & Poor Fen) are relatively rare and in difficult to access locations, and few degraded examples are available.

TABLE B6. Site Numbers available from natural heritage programs based on assigning their wetland records to macrogroups and condition strata.

	Condition Stratum				
Macrogroup Stratum	**A**	**B**	**C**	**D**	**Total**
Northern & Central Floodplain Forest	7	14	14	4	39
Northern & Central Swamp Forest	1	16	12	4	33
Appalachian, Interior Plateau & Prairie Fen	1	26	13	4	44
North American Boreal Bog & Fen	1	16	13		30
Atlantic & Gulf Coastal Plain Pondshore & Wet Meadow	7	24	7		38
Eastern North American Wet Shrub, Meadow & Marsh	3	28	15	9	55
Great Plains Freshwater Wet Meadow, Wet Prairie & Marsh	3	15	13	7	38
Total	23	139	87	28	277

Samples for the original design were chosen from the suite of existing sites randomly, using the "Random Selection within Subsets" tool in the Hawths Tools extension for ArcGIS (Beyer 2004). This tool picks a random subset of features from all features in a GIS shapefile using user-defined strata. The selection of samples for one macrogroup is illustrated on Figure B4. The resulting samples were reviewed to ensure adequate representation across ecoregions. Between the state EO data and MiRAM data, we achieved over 95% of our target number of 280 points using known sites. Two thirds of the data were to be collected in Michigan, one third in Indiana. The remaining 5% of sites were attained by relying on the expertise of the field staff, who could suggest potential sites where missing combinations of macrogroup and condition were needed.

FIGURE B4. Location of all sites and their EORANK for a macrogroup (Rich Fen) based on state program databases, and categorized by whether or not they were selected as part of the sample design.



In order to maximize the amount of real field time available and spend less time in transit, we preferentially sampled sites located in somewhat close proximity to one another. This would speed up field work considerably. However, in order to avoid spatial autocorrelation, we maintained a minimum separation distance of 0.5 km for sites within the same macrogroup. This distance is consistent with the surrounding landscape area used for the calculation of Landscape Context metrics (Table B4). In practice, the vast majority of field sites within the same macrogroup were located at least 1 km apart.

An overdraw of sites was made to account for the possible need to replace sites from this original random draw. The overdraw pool included one additionally randomly selected site per strata. Because crews were state-based, separate overdraw pools were created for Indiana and Michigan. Additional replacement sites could be selected later as necessary, by choosing randomly from the remaining pool of known sites for that stratum.

We gave some consideration in our site selection to issues of public/private ownership. Our random sample draw methodology described above did not explicitly take into account land ownership or favor clustered samples. However, once a provisional selection of sites had been made, we allowed field crews to swap sites based on accessibility. Publicly owned sites required less preparation, because access permission is more readily obtained. Over 90% of the original sites selected were part of the final sites sampled.

It was apparent that our overall pool of sites was relatively short on both A-ranked and D-ranked sites, so crews were instructed to identify additional occurrences of such sites during the course of their survey work, and determine if they met the needs of the sample design.

Caveats

Although designed to minimize bias, the methodology is not unbiased. There are geographic or ownership biases that underlie sites catalogued by natural heritage programs. Consequently, we cannot be certain that the suite of sites from which we chose our project sample represents the true population distribution of reference sites. For those combinations of wetland type and condition with fewer available natural heritage sites (such as D-ranked sites), most or all existing sites were included in the project sample. Thus, even as we attempted to maintain a geographic spread, any spatial and ownership biases within the program and MiRAM datasets were carried into our data set. Likewise, supplementing the randomly chosen sample with additional sites identified based on drive-by selection or local expert knowledge may have introduced additional biases into our sampling methodology.

Despite these violations of truly random and unbiased sampling, we are confident the sample design satisfied the goals of the project—to ensure that the full range of variation in wetland types and conditions are sampled across the study area in an efficient manner. Given data limitations present at this time, random selection methods were used to the greatest degree possible, and our method allows inference with regard to the range of conditions of all major wetland types within the project area.

Statistical Tests of Sampling Design

We evaluated our sampling design by examining the distribution of actual condition ratings (from the field results of our ecological integrity assessments, as summarized in Section C) against the predicted condition ratings from our condition stratum methods. Ideally, we would like to ask how well these ratings might predict overall ecological integrity (based on landscape context, condition, and size), but comparisons with the overall IEI would be misleading because we used the same Landscape Context ratings when calculating an overall IEI as we did for the condition stratum. These measures are correlated by definition. So instead our main focus was on how well we predicted on-site Condition and Vegetation. Our reference information was the EORANK, which played no role in assigning our Ecological Integrity rating.

We chose a one-way analysis of variance (ANOVA) (appropriate to categorical rating data) to see how well the condition stratum performed. We used the **Kruskal-Wallis rank sum test,** which is the non-parametric analogue of a one-way ANOVA (when there is one nominal variable and one measurement variable and the measurement variable does not meet the normality assumption of an ANOVA). The Kruskal-Wallis test does not make assumptions about normality. Like most non-parametric tests, it is performed on ranked data, so the measurement observations are converted to their ranks in the overall data set: the smallest value gets a rank of 1, the next smallest gets a rank of 2, and so forth. We applied the **Pairwise Wilcoxon Rank Sum Tests** to calculate pair-wise comparisons between condition ratings with corrections for multiple testing, making adjustments to p values when testing multiple comparisons. Finally, comparisons were scanned using **Notched Boxplots.** If the two boxes' notches do not overlap, this is "strong evidence" that their medians differ (Chambers et al, 1983, p. 62).

B.3. RESULTS

Reference Gradient by Classification Stratum

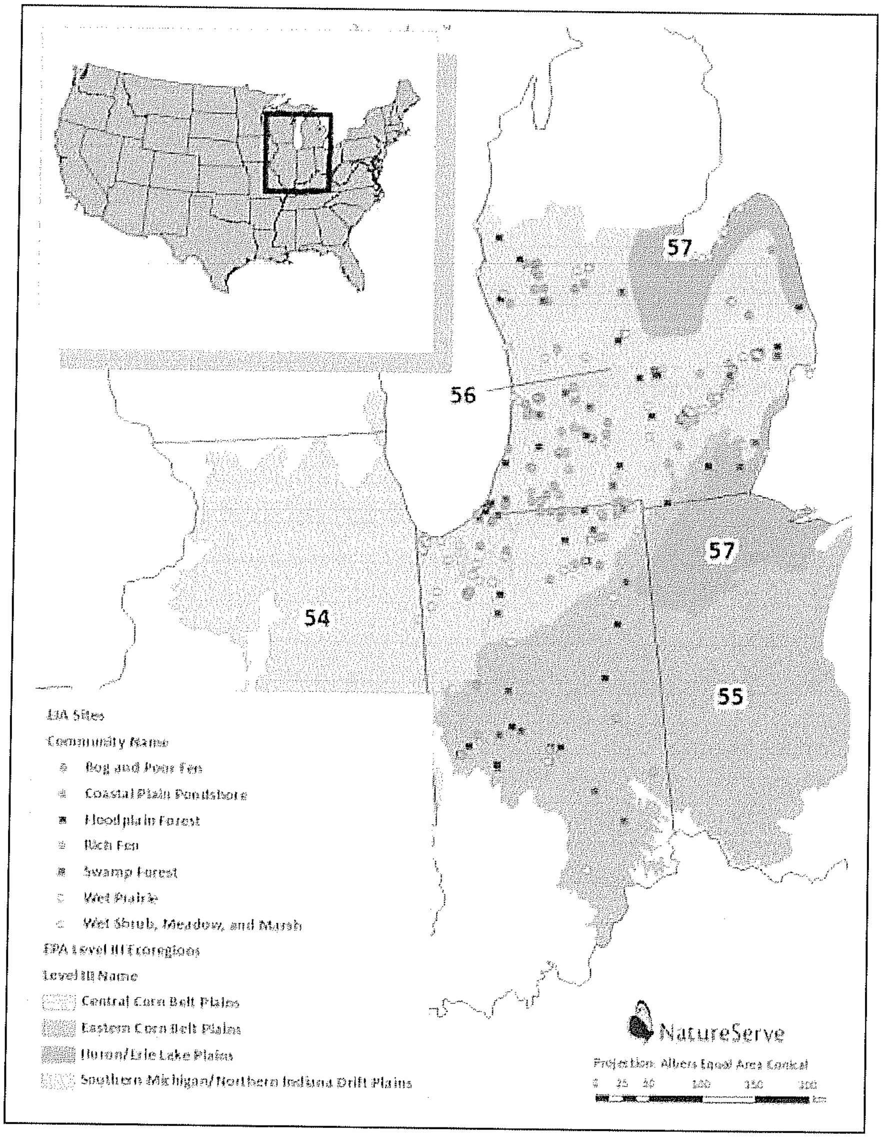
Macrogroups and State Types

Our sampling design resulted in a range of wetland sites spread across the 7 macrogroups found in the study area. We maintained a fairly consistent level of sampling across all macrogroups (30–55 sites per macrogroup), spread across the ecoregions (Table B7, Fig. B5).

TABLE B7. List of Formations and Macrogroups found in the study area, and number of study sites found for each macrogroup. Colloquial names for macrogroups are provided in square brackets.

NVC FORMATION **Macrogroup**	**Number of Sites**
SWAMP & FLOODED FOREST	
Northern & Central Floodplain Forest (M029) **[Floodplain Forest]**	39
Northern & Central Swamp Forest (M030) **[Swamp Forest]**	33
FRESHWATER SHRUBLAND, WET MEADOW & MARSH	
Eastern North American Wet Shrub, Meadow & Marsh (M069) **[Wet Shrub, Meadow & Marsh]**	55
Great Plains Wet Meadow, Wet Prairie & Marsh (M071) **[Wet Prairie]**	38
Atlantic & Gulf Coastal Plain Pondshore & Wet Meadow (M067) **[Coastal Plain Pondshore]**	38
BOG & FEN	
Appalachian, Interior Plateau & Prairie Fen (M061) **[Rich Fen]**	44
North American Boreal Bog & Fen (M062) **[Bog & Poor Fen]**	30
Total	277

FIGURE B5. Sites selected for the reference gradient of wetland types.



Within each macrogroup, we also sampled all state natural community wetland types at least once (Table B8).

TABLE B8. State Natural Community Types by NVC Macrogroup.

Macrogroup State & State Natural Community Type	Number of Occurrences
Northern & Central Floodplain Forest (total)	39
IN (total)	17
Mesic Floodplain Forest	8
Wet Floodplain Forest	2
Wet-mesic Floodplain Forest	7
MI (total)	22
Floodplain Forest	22
Northern & Central Swamp Forest (total)	33
IN (total)	14
Bluegrass Till Plain Flatwoods	1
Boreal Flatwoods	1
Central Till Plain Flatwoods	5
Forested Fen	3
Sand Flatwoods	2
Swamp Forest	2
MI (total)	19
Hardwood-Conifer Swamp	1
Rich Conifer Swamp	3
Rich Tamarack Swamp	5
Southern Hardwood Swamp	7
Wet-mesic Flatwoods	3
North American Boreal Bog & Fen (total)	30
IN (total)	10
Acid Bog	10
MI (total)	20
Bog	20
Great Plains Freshwater Wet Meadow, Wet Prairie & Marsh (total)	38
IL (total)	1
Wet Prairie	1
IN (total)	9
Wet Prairie	2
Wet Sand Prairie	2
Wet-mesic Sand Prairie	5
MI (total)	28
Interdunal Wetland	5
Lakeplain Wet Prairie	6
Lakeplain Wet-mesic Prairie	5
Wet Prairie	5

Wet-mesic Prairie	4
Wet-mesic Sand Prairie	3
Appalachian, Interior Plateau & Prairie Fen (total)	44
IN (total)	15
Circumneutral Bog	3
Circumneutral Seep	3
Fen	7
Marl beach	1
Panne	1
MI (total)	29
Prairie Fen	29
Eastern North American Wet Shrub, Meadow & Marsh (total)	55
IN (total)	24
Acid seep	1
Marsh	11
Sedge Meadow	4
Shrub Swamp	8
MI (total)	31
Emergent Marsh	6
Great Lakes Marsh	5
Inland Salt Marsh	3
Inundated Shrub Swamp	1
Northern Wet Meadow	1
Southern Shrub-carr	3
Southern Wet Meadow	12
Atlantic & Gulf Coastal Plain Pondshore & Wet Meadow (total)	38
IN (total)	12
Muck Flat	9
Sand Flat	3
MI (total)	26
Coastal Plain Marsh	26
Grand Total	277

Hydrogeomorphic Classification

For each wetland occurrence visited, we assigned the HGM class (Table B9). The kind and number of HGM classes found within each macrogroup is reported in Table B10. The most predominant HGM type in the study was Depressional. Organic Flats are relatively rare in these ecoregions.

TABLE B9. HGM type and number of occurrences for the project. "Primary only" refers to it being listed as the only HGM class at a site.

HGM Class	Primary Only	Primary + Secondary	Types of Secondary Classes
Depressional	118	6	Mineral Soil Flats (3), Organic Soil Flats (1), Riverine (2)
Lacustrine Fringe	24		
Mineral Soil Flats	23	1	Organic Soil Flats
Organic Soil Flats	11	2	Slope
Riverine	48	1	Slope
Slope	43		
Grand Total	267	10	

TABLE B10. The variation in HGM primary class within each NVC macrogroup.

Macrogroup: HGM type	Number of Sites
Northern & Central Floodplain Forest	**39**
Mineral Soil Flats	1
Riverine	37
Slope	1
Northern & Central Swamp Forest	**42**
Depressional	24
Lacustrine Fringe	2
Mineral Soil Flats	8
Organic Soil Flats	1
Riverine	1
Slope	6
Atlantic & Gulf Coastal Plain Pondshore & Wet Meadow	**38**
Depressional	30
Lacustrine Fringe	2
Mineral Soil Flats	4
Organic Soil Flats	2
Eastern North American Wet Shrub, Meadow & Marsh	**46**
Depressional	25
Lacustrine Fringe	11
Mineral Soil Flats	1
Organic Soil Flats	1
Riverine	5
Slope	3
Great Plains Wet Meadow, Wet Prairie & Marsh	**38**
Depressional	19

Lacustrine Fringe	4
Mineral Soil Flats	10
Riverine	4
Slope	1
Appalachian & Interior Plateau Bog & Fen	**44**
Depressional	3
Lacustrine Fringe	4
Organic Soil Flats	2
Riverine	3
Slope	32
North American Boreal Bog & Fen	**30**
Depressional	22
Lacustrine Fringe	1
Organic Soil Flats	7
Grand Total	**277**

Reference Gradient by Condition Stratum

Field crews collected data that permitted us to calculate the Ecological Integrity rating for each wetland, as well as the on-side Condition and Vegetation scores (see Faber-Langendoen et al. 2011c). We can thus examine how well the overall predicted set of wetland conditions, based on the "condition stratum" matched the final measured condition and vegetation ratings. Table B11 shows the number of sites for each condition rating based on field measures, as well as the predicted rating based on the individual or combination of factors used for the Condition stratum—Landscape Context alone, EORANK alone, and various combinations of the two (see *Methods: Classification and Condition Strata* above).

TABLE B11. Actual and Predicted A–D Condition ratings based on various condition stratum factors, alone or in combination. Not shown are the macrogroups.

Condition Stratum	**A**	**B**	**C**	**D**	**Un-as-signed**	**Grand Total**
Actual A–D condition ratings from field	127	126	23	1		277
Landscape Context	57	122	73	25		277
EORANK	55	78	73	20	51	277
Condition Stratum 1 (Landscape Context X EORANK – rigorous A)	23	139	87	28		277
Condition Stratum 2 (Landscape Context X EORANK – moderate A)	66	135	59	17		277

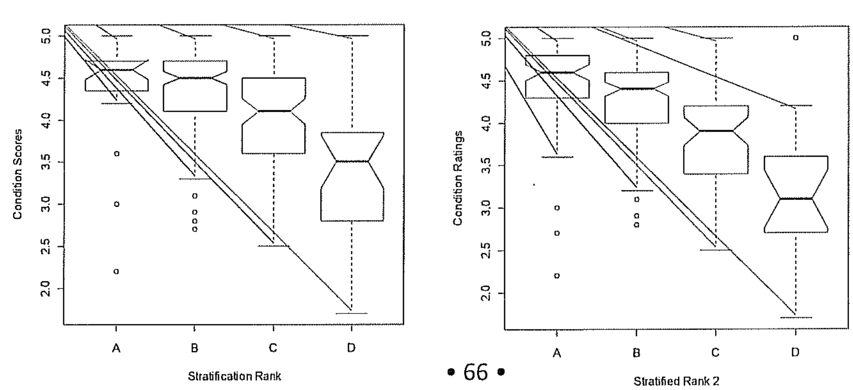
Our key question then is: Did our condition stratum approach ensure that we sampled a full range of wetland condition, based on both our original Condition Stratum approach 1 and Condition Stratum 2? Our results are summarized in Table B12 and Figure B6 (see also Table B11). Looking first at Condition (Vegetation, Hydrology and Soils), we can see that on-site Condition is best predicted by the combination of Landscape Context and EORANK, using either approach 1 or 2, but our 2nd approach is an improvement (in which our requirements for an A rating were not as rigorous) (Fig. B6). Thus, our first condition stratum approach underestimated the number of A-rated sites compared to our second approach.

Our condition stratum 2 approach was also the best predictor of Vegetation ratings (Table B12). Vegetation ratings are also less influenced by landscape context, and more informed by knowledge of on-site condition, as provided through the EORANKs.

TABLE B12. F-values for one-way Kruskal-Wallis rank sum test on the various stratification approaches: Landscape Context alone, EORANK alone, Stratification Rank 1 (Landscape Context X EORANK stringent A requirements) and Stratification Rank 2 (Landscape Context X EORANK moderate A requirements). All F values have $p < 0.001$, indicating all factors successfully distinguish A/B from C from D. See also Figure B6.

Factors	Condition Score	Vegetation Score
Landscape Context	52.0	25.1
EORANK	47.7	35.0
Condition Stratum 1	54.8	33.1
Condition Stratum 2	68.8	40.8

FIGURE B6. Comparison of predicted on-site Condition scores based on the Condition stratum method 1 (Stratification Rank) and Condition stratum method (Stratified Rank 2).



B.4. Discussion

Classification

In creating a sampling design for our reference gradient, our first concern was to sample the full range of wetland types. As shown in Table 8, we were able to do this by linking the fine grained state natural community types to the mid-level NVC macrogroup. We also sampled the full range of HGM types found in the region: we had a preponderance of Depressional types, and few Organic Flats types. Thus, using a classification stratum based on NVC macrogroups ensured that our sampling covered a wide range of wetland types. Because we could confidently assign sites to macrogroups, we were rather successful in maintaining a balanced set of sites across macrogroups (30–55 sites per macrogroup).

Condition

Our second concern for our reference gradient was to sample the full range of wetland conditions. Our approach was to rely on a combination of factors to establish a condition stratum—first, the readily observable landscape context factors available from imagery, and, second, the expert-based field evaluations of on-site conditions (especially vegetation) recorded by state natural heritage ecologists. We found that, by using both criteria, we were more successful at predicting the full range of reference gradient conditions than either landscape context or on-site rank alone (Table 12). Our field design was executed using condition stratum 1, which under predicted the number of A-ranked sites we might encounter and over predicted the number of D-ranked sites. Thus we sampled fewer D-ranked sites than were predicted by that approach. We recommend using our revised condition stratum design for future studies.

Others have also found that landscape context metrics alone, based on remote sensing imagery, have only limited value in predicting on-site condition (Mack 2006, 2007, Mita et al. 2007). This suggests caution in using landscape alone as a predictor of individual site conditions, though it is helpful for assessing overall watershed condition (e.g., Tiner 2004).

The stratification methods we use here bode well for identifying a reference gradient of wetlands across many parts of the country. Program data are widely available across the country. Indeed, we have recently compiled all natural heritage data into a master database, in which macrogroups have been assigned to all state records. Currently available remote sensing imagery can be used to calculate the landscape context metrics, and these can be combined with the condition stratum method 2 to provide a robust prediction of on-site wetland condition. These data can be a primary source of reference sites for studies needing a reference gradient for wetlands, grasslands, forests and other types. In fact, EPA's National Wetland Condition Assessment is currently using these data as part of their process to identify approximately 150 benchmark reference standard wetland sites across the country to help inform their wetland condition assessment (Faber-Langendoen et al. in prep, G. Serenbetz pers comm. 2011). Knowledge of these sites is becoming increasingly important, given continuing levels of conversion or degradation of native ecosystems across many parts of the country.

SECTION C: TESTING THE ECOLOGICAL INTEGRITY ASSESSMENT METHOD

C.1. INTRODUCTION

There is a growing body of wetland assessment methods that provide standardized field sampling and reporting methods for assessing wetland condition (e.g., Mack 2001, 2004, Collins et al. 2007, Jacobs et al. 2010, see summary in Fennessy et al. 2007). Data on the ecological condition of wetlands can be used for ambient monitoring of wetland status and trends, to prioritize sites for conservation or restoration, guide mitigation applications at site and watershed scales (Faber-Langendoen et al. 2008) and contribute to land use planning. Much has been done to develop and test methods for assessing wetland condition, including both rapid and intensive methods (Mack 2001, 2004, 2006, Collins et al. 2007, Miller et al. 2006, 2007, Fennessy et al. 2007). But many studies have been local in geographic scope, or restricted to a subset of wetlands in a region, or have not provided a larger framework within which to assess changes in wetland condition. There is a need to provide methods that assess condition across the full range of wetland types and condition in a region and provide summary reports that make the results accessible to a wide range of audiences.

Our purpose in this section is to apply our ecological integrity assessment method to the 277 wetland sites that span the reference gradient of wetland types and conditions across the study areas (as identified in Section B above). We then completed a "post-hoc" analysis of the indicator data collected across the sites to verify that the metrics and scores can discriminate among a range of conditions, from "excellent" (minimally disturbed) to "poor" (degraded) wetlands. We also created a scorecard and index of ecological integrity (IEI) for reporting on wetland integrity.

C.2. METHODS

Our site selection methods for the 277 sites are fully described in Section B (see Table B6, Fig. B5 for summaries). Here we describe the ecological integrity assessment field methods used to survey each site.

Level 2: Rapid Field Assessment Methods

Overview of Field Methods

A field manual was developed to guide the use of field forms by the crews (Faber-Langendoen 2011). The general procedure for conducting a Level 2 assessment consisted of a series of steps (adapted from Collins et al. 2006, Chapter 3):

PRELIMINARY SITE SELECTION (OFFICE)

Step 1: We reviewed the occurrence list for the field season, based on the sampling design established for the project, including the primary sites and backup sites. Site selection was determined by the wetland types and their conditions. The statistical design of the study was set up to avoid sampling two wetlands in close proximity that were also of the same type (because the area sampled as part of the landscape context would overlap, making the samples within a type non-independent).

Step 2: For each wetland occurrence at a site, we assembled background information about the ownership, access, size, condition (=EORANK), spatial location (many sites had the spatial extent of the wetland type mapped in GIS), and state wetland classification type. We also assembled aerial photo imagery for the site. We made landowner contacts, as needed, before accessing the site.

Step 3: We reviewed the classification of the wetland, starting from the state program classification, which was nested within the U.S. National Vegetation Classification at the macrogroup level (See Table B8 above). Additional classifications include NatureServe's Ecological Systems (Table B3 above) and the Hydrogeomorphic (HGM) classification (Tables B9, B10). Field crews needed to identify the state natural community type and the HGM class.

State Natural Community Type:

a) Michigan. See Kost et al. (2007) and *web4.msue.msu.edu/mnfi/*

b) Indiana. See Jacquart et al. (2002) and *www.in.gov/dnr/naturepreserve/4743.htm*

Descriptions for the macrogroups are being compiled and will become available on the NatureServe website at *www.natureserve.org*, and the USNVC partners website at *usnvc.org.*

Step 4: We verified the appropriate season and other timing aspects of field assessment for sampling various wetland types.

SITE SELECTION REVIEW (OFFICE)

Step 5: We defined the assessment area (AA) as an area of given condition and wetland type (at state natural community scale) and small enough in size to be observable in the course of a 2–4 hour visit (Rocchio 2007). Accordingly we define the AA as "the entire area, subarea, or point of an occurrence of a wetland type with a relatively homogeneous ecology and condition." Practically speaking, this meant AAs had to be less than 20 ha (50 acres). For large wetland occurrences (> 20 ha), we determined which portion of the wetland could be visited. AAs could not always be determined prior to the field visit, so adjustments were made in the field.

Although 20 hectares is too large to survey intensively, the crews made a judgment as to whether the area they surveyed appeared typical of the entire AA or the polygon or EO within which the AA occurred. The advantage of this approach is that a "polygon" or "wetland occurrence" focus is maintained, rather than a "point-based" approach (See Fennessy et al. 2007, Faber-Langendoen et al. 2011a). For many applications, the goals of the EIA are to determine the condition of an extensive area of a wetland occurrence or polygon.

In some rapid assessments, the type and condition are ignored and the entire wetland is assessed as part of the AA. In other assessments, detailed guidelines for establishing AAs are provided (ORAM, CRAM). Our methodology follows the latter approach, consistent with natural heritage methodology, where an occurrence of a wetland type of conservation or management significance is tracked based on its type, size and relatively uniform condition.

LANDSCAPE CONTEXT EVALUATION (OFFICE)

Landscape Context metrics (L1) are extrapolated from remote sensing imagery. The goal is to develop metrics that assess the landscape context (and thereby on-site conditions of an ecosystem). Satellite imagery and aerial photos are the most common sources of information for these assessments. Typically it is the stressors to the ecological integrity that are most observable with these sources of information, so condition is generally inferred from stressors.

Step 6: Assess Landscape Context

Primary Approach: Remote Sensing Landscape Metrics. NatureServe staff, using both satellite imagery and aerial photography, established the buffer (200 m) around the Assessment Area (AA) polygons, and, using the centroid of the polygon, established the circular areas that comprise the "core" (100 ha) and "supporting landscape" contexts of the AA (1000 ha)[10]. We analyzed the imagery to calculate the scores and ratings for the core and supporting landscapes and buffer metrics for each occurrence at a site. The spatial boundaries of the landscapes and buffer and the metric scores were moved into the database.

Secondary Approach: Landscape Condition Model. NatureServe has developed a Landscape Condition Model (LCM, Comer and Hak 2009), similar to the Landscape Development Index used by Mack (2006). The model provides a single stressor-based index that integrates the effect of multiple landscape stressors on overall landscape condition. The algorithm for the model uses 30 m resolution pixels from various land use layers (roads, land cover, water diversions, groundwater wells, dams, mines, etc.). These layers are the basis for various stressor-based metrics. The metrics are weighted according to their perceived impact on ecological integrity, into a distance-based, decay function to determine what effect these stressors have on ecological integrity. The result is that each grid cell (30 m) is assigned a "score." The product is either a watershed or landscape map depicting areas according to their potential "integrity," or the condition of individual polygons or patches can be characterized. The index is segmented into three or four rank classes, from Excellent (minimally disturbed) (A) to Poor (degraded) (D) (See Appendix 2 for more details).

ON-SITE CONDITIONS (FIELD)

Step 7: For the field visit, standard field forms were used (Faber-Langendoen 2011). Further details on the field methods are presented below.

7a. All sites were assessed using the rapid assessment (L2) method, including basic description (vegetation and environmental characteristics), integrity metrics and stressor evaluation.

7b. One-third of the sites were further assessed using the intensive assessment (L3) method, including a 0.1 ha plot. Plots focused on the vegetation, recording all species and their cover, and recording stem diameters and density for all tree stems $\geq$ 10 cm dbh, along with basic soil and hydrology information to help characterize the wetland type.

DATA MANAGEMENT (POST FIELD)

Step 8: All data were entered into an Ecology Observations Database (Access ©).

---

[10] In future versions of our protocol, we intend to use a "buffered polygon or point" approach, defining an inner buffer of 100 m, the core landscape of 250 m, and the supporting landscape of 500 m. This provides a more consistent landscape context assessment protocol for both buffer and landscape metrics.

8a. Data clean-up was completed, including determination of appropriate plant species taxonomy and ensuring the GPS coordinates accurately represent the location of the AA, etc.

8b. Data entry into NatureServe's Ecological Observations database was completed. Data in 2009 were stored separately for IN and MI. In 2010, several changes were made to the protocol, particularly for stressors, necessitating a slightly different design. The 2010 data from both states were managed in a single database. The database provided a scorecard for each wetland and export formats for analyses.

8c. QA/QC procedures were completed by the state program data entry staff and NatureServe data management staff.

Step 9: Core data were uploaded into the state program databases to upgrade site information, classification, and EORANK.

Level 2 Field Protocols

A field crew (usually two people) typically conducts a rapid (L2) field assessment within two to three hours, plus two hours preparation time assessing the imagery. Once the crew leaves the field, the field forms are essentially complete, apart from data cleanup and QA/QC. Additional time may be needed on plant species taxonomy issues, and to ensure that the GPS coordinates accurately represent the location of the AA.

All field crews had at least one person who was trained in ecology, with sufficient botanical expertise to recognize the major elements of the flora. Crews also had general experience with hydrology and soils, sufficient for the rapid assessment methods. One-day field training exercises were conducted in early May of each season in order to ensure consistent application of the field protocols.

Upon arriving at a site, the crews were asked to validate the classification of the wetland community to the state type, and thereby the NVC macrogroup (Table B8 above). If changes were needed, crews documented these changes. Crews also assigned types to HGM class.

Typically, crews had pre-existing polygons to visit, based on natural heritage element occurrence maps. Crews checked the polygon boundaries of the map to verify or update the extent of the occurrence as part of the field survey. Readily observable ecological criteria such as vegetation, soil, and hydrological characteristics are used to define wetland boundaries, regardless of whether they meet jurisdictional criteria for wetlands regulated under the Clean Water Act.

If the wetland was small enough to survey in its entirety, then the AA and the EO boundaries were synonymous But if the EO was very large or if an EO had had two or more conditions present), then the AA was restricted to a portion of the occurrence. Notes on the AA boundaries were made using GPS and hand-drawn field notes on aerial photos and maps.

Level 2 Metrics and Field Forms

Standard field forms were used by all field crews in both states (Faber-Langendoen et al. 2011d). The list of metrics collected by the field crews is shown in Table C2. Some metrics vary by NVC formation/macrogroup and HGM class, so crews used their assignment of the state to the state type or HGM class to guide their use of metrics. Field forms are structured by major categories: GENERAL DESCRIPTION (Site, Location, and Classification), DETAILED DESCRIPTION (optional narrative form), VEGETATION PROFILE, ENVIRONMENTAL PROFILE, ECOLOGICAL INTEGRITY METRICS (Vegetation, Hydrology, Soil, Size, Buffer), STRESSOR METRICS (Vegetation, Hydrology, Soil, Buffer).

TABLE C2. List of Metrics collected for Level 2 Ecological Integrity Assessments. Protocols for each metric are provided in Appendixes. Metrics in italics were later dropped after statistical assessment (see Table C10).

RANK FACTORS	MAJOR ECOLOGICAL FACTORS	METRICS
LANDSCAPE CONTEXT	LANDSCAPE	Connectivity
		Land Use Index
	BUFFER	Buffer Index
SIZE	SIZE	Relative Patch Size
		Absolute Patch Size (ha)
CONDITION	VEGETATION	Vegetation Structure
		Organic Matter
		Regeneration (woody)
		Native Plant Species Cover
		Invasive Plant Spp. – Cover
		Increasers – Cover
		Vegetation Composition
	HYDROLOGY	Water Source
		Hydroperiod
		Hydrologic Connectivity
	SOIL	Physical Patch Types
		Soil Disturbance
		Water Quality

To survey the AA at the L2 level, field crews walked through and visually assessed the site, and recorded the measures specified on the field form. [11]

Level 2 Stressors Checklist

Stressor information can be a useful addition when evaluating the ecological integrity of an occurrence (Kapos et al. 2002). Typically, they aid in further understanding the overall condition of a wetland. Stressors were recorded by Major Ecological Factor (Buffer, Vegetation, Soils, Hydrology). Stressors were listed if and only if they were observed or inferred to be occurring, and were not included if they were projected to occur, but do not yet occur. Stressors were characterized in terms of scope and severity. **Scope** is defined as the proportion of the occurrence that can reasonably be expected to be affected by the stressor. Within that affected scope, **severity** is the intensity of damage to the occurrence that can reasonably be expected from the stressor. Stressor scope and severity are used to create field-based versions of stressor indices. Individual stressors as rated as Very High, High, Medium, Low, based on the combination of Scope and Severity (see Appendix 3: Level 2 Protocols – Stressors Checklists).

Standardized stressor checklists have been developed for a variety of rapid assessment methods (Collins et al. 2006, Faber-Langendoen et al. 2008, Collins and Fennessy 2010). Our list was adapted from those sources (Faber-Langendoen 2011).

Level 3: Intensive Field Assessment Methods

Integration with Level 2 Methods

Our Level 3 method builds on the Level 2 methods. Of the 277 sites that are visited, 88 sites also had a Level 3 assessment completed. The level 3 data are restricted to a single representative area —a 0.1 ha plot—within the AA, whereas the level 2 assessment is completed on the entire Assessment Area. Notwithstanding these differences in scale, we compared the level 3 assessment data with the level 2 assessment data.

Vegetation Sampling Methods

We used a vegetation sampling protocol that provides information on basic vegetation composition and structure and was suitable for extracting metrics for the Floristic Quality Index (FQI).

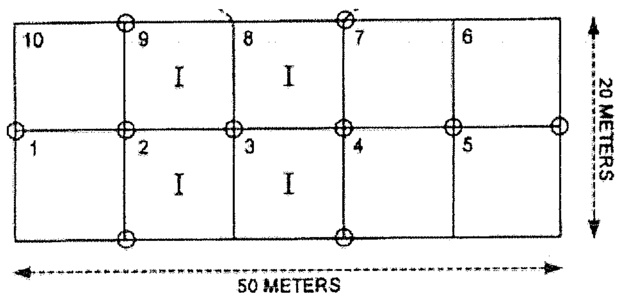
PLOT METHOD

A 20 m x 50 m reléve plot was used to collect vegetation data. The method is based on the historic Whittaker diversity plot method, modified by Peet et al. (1998), and used by other wetland

[11] Although not used, here, it may be desirable in some circumstances to add a "quick" vegetation plot to the level 2 evaluation. For example, a 50 m tape is laid, with 10 m flags on each outer end, and vascular plant species presence and cover are recorded by strata by surveying the 10 m wide area on each side of the tape). A few key structural attributes could also be collected (e.g., coarse woody debris, number of stems > 30 cm dbh. The plot can be subjectively or objectively placed within the AA to represent the typical heterogeneity within the AA. Such a plot would provide valuable quantitative information on vegetation cover, species richness and abundance to supplement the rapid assessment, but pushes the methodology towards a hybrid of Level 2 and Level 3.

assessment methods (Mack 2004, 2007a,b, Rocchio (2007). The structure of the plot consists of ten 100 m^2 modules (a total of 1000 m^2 or 0.1 hectare) which are typically arranged in a 20 m x 50 m array (Fig. C1).

FIGURE C1. Relève Plot Method (from Peet et al. 1998). I = intensive modules 2, 3, 8, 9.



Plots were laid out using a 50 m measuring tape, extended as the centerline of the plot from an origin. Starting at the origin (zero), a stake flag (or flagging tied to vegetation) was placed every 10 m. Red stake flags or flagging were placed at the 0, 40, and 50 m marks and green stake flags/flagging at the 10, 20 and 30 m marks. This helped visualize the four "intensive modules" which occur on either side of the centerline between the 10–30 m marks. Next, a 10 m rope was extended perpendicular on either side of the centerline at each 10 m mark. Red or green flags were placed at the end of the rope to mark the lateral boundaries of each module and the plot.

The plots were located subjectively by the field crews. Under typical conditions, the specific location was chosen because it was judged to contain structure(s) and composition(s) typical of the observed wetland, or to contain the most frequently occurring structure(s) and composition(s). If the wetland had an irregular shape and 20 m by 50 m plot would not "fit" into the specified wetland, the 2 x 5 array of modules were restructured to accommodate the shape of the wetland or AA. For example, a 1 x 5 array of 100 m^2 modules was used for narrow, linear areas. A 2 x 2 array of 100 m^2 modules was used for small sites (Peet et. al. 1998; Mack 2004). Regardless of the structure, a minimum of four intensive modules were always sampled.

Each module in the plot was numbered by standing at the 0 m mark facing the 50 m end (Figure 5). The modules were assigned from 1–5 starting on the right side and modules 6–10 were assigned using a similar method then from the 50 m mark. Intensive modules were typically number 2, 3, 8, and 9. For those plots that did not use a 2x5 array of modules (e.g. 1x5 or 2x2), the module numbers may be different (and were randomly chosen).

Floristic measurements include presence/absence, first made within the four core (or "intensive") 100 m^2 modules (2, 3, 8, 9). Crews recorded all species in the first module, then added any new species module by module for the remaining three modules. They evaluated percent canopy cover of each species across all four modules (400 m^2). They then searched the remaining six modules ("residuals"), adding any new species and cover values for those not found in the intensive modules (see Yorks and Dabydeen 1998 on the importance of this additional search area). They also noted any species of interest found within the wetland AA but not in the plot. Thus the final values obtained were a single species list with cover values across a 0.1 ha area. Tree species cover was assessed separately for seedlings, saplings, and overstory. Overstory tree cover was assessed across the entire 0.1 ha plot regardless of what modules they occurred in.

Cover was visually estimated at the level of the 100 m^2 module (depth 1) using the following cover classes (Peet et al. 1998): 1 = trace (one very small individual), 2 = 0.1–<1%, 3 = 1–<2%, 4 = 2–<5%, 5 = 5–<10%, 6 = 10–<25%, 7 = 25–<50%, 8 = 50–<75%, 9 = 75–<95%, 10 = > 95%.

Tree Structure

LIVE STEMS

Level 3 vegetation sampling also included a stem profile. Information on size and number of tree stems were collected by tallying tree stems (separately by species) to the nearest 10 cm intervals from (1) 10–49 cm dbh, and recording the diameter of each stem to the nearest cm for stems greater than 50 cm dbh. Stem information was collected on all stems > 10 cm dbh. Measuring the diameter-breast-height (dbh) for each stem in the plot allows calculation of basal area, a widely used measure of tree abundance, and density. Measures of dbh can also provide important information on stand dynamics and structure not captured by cover estimates. This cut-off can be lowered in environments where mature trees may often be much smaller or in cases where information on the regeneration (sapling and/or seedling) layers is needed.

STANDING SNAGS

All standing snags (dead standing tree boles >10 cm and at least 1.4 m tall) were recorded.

FALLEN LOGS (COARSE WOODY DEBRIS)

Fallen logs were defined as dead fallen tree trunks greater than 10 cm in diameter. Each fallen log was recorded by diameter and length. The diameter was assigned using size class categories from 10–50 cm, and to nearest cm for over 50 cm. The length was assigned to nearest m for that part of the stem that is within the plot and which exceeds 10 cm diameter.

Soil Sampling Methods

In addition, soil and substrate characteristics were collected at each site to aid in characterization of the wetland at the site. A soil core or auger was used to estimate soil values in each of two locations. A maximum depth of 50 cm was sufficient to record the following: Depth to Impervious Layer, Depth to Saturated Soils, Depth to Water Table, Organic Soil (Sapric (muck), Hemic (mucky peat), or Fibric (fibric peat)), Mineral Soil Texture, and any additional comments were recorded about the soils (e.g., presence of marl layers, irregular depressions, mounds, etc.).

Sample Handling and Processing

Plant specimen data were handled in typical fashion (see Mack 2007b for details). Standard state nomenclature was used, and then standardized to PLANTS / NatureServe taxonomy for vascular plant, nonvascular plant, and lichen names as accepted by NatureServe's standard references, which represent the consensus standards for researchers working in a given geographic area (see Kartesz 1999).

We coordinated this standard nomenclature with the standard state nomenclature used in Michigan and Indiana.

Data Management

Level 2 Data Management

All data were entered and managed in an Ecological Observations Database that was specifically designed for the project, yet structured as generically as possible to provide an ongoing database tool for other ecological integrity assessment projects. The database is structured to match field data protocols: General Site Description, Level 2 metrics, Level 2 stressor checklists, and Level 3 metrics, including vegetation plot data. Data in 2009 were stored separately for IN and MI. In 2010, several changes were made to the protocol, particularly for stressors checklists, necessitating a slightly different design. The 2010 data from both states were managed in a single database. An Index of Ecological Integrity (IEI), including a scorecard, was used within the database to summarize all metric ratings for L2 assessments (see *Index of Ecological Integrity* section below). Data are available from NatureServe and from the network programs upon request.

Level 3 Vegetation Data management

Vegetation data were entered into a Microsoft Access™ database, where cover class data were transformed into cover values (the midpoint of each cover class). Mean cover for each species were averaged across the intensive modules and used in data analysis. For those species only occurring in the residual plots, the cover value for the residual plots was used for analysis. To eliminate spelling errors, a drop-down list was used for species entry. Unknown or ambiguous species (e.g. *Carex* sp.) were recorded but not included in data analysis. Data entry was reviewed by an independent observer for quality control.

The Michigan Floristic Quality Assessment (FQA) database (Herman et al. 2001) was used to populate life history traits, wetland indicator status, and C-values in the data reduction spreadsheet for each species in the plot. Species nomenclature follows USDA PLANTS Database *http://plants.usda.gov/*) as of January 2009. Since many practitioners in Michigan use the Michigan Flora by Voss (1972, 1985, 1996), and in Indiana use a variety of floras as a field key and nomenclature reference, these names were cross-referenced to the PLANTS names in the FQA database.

Index of Ecological Integrity

IEI and Scorecard

As described above, our methodology included a prototype Index of Ecological Integrity (IEI) for Level 2 assessments. Our selection of metrics for the IEI was based on reviewing tests of these metrics from the literature (e.g. for ORAM see Mack 2006, 2007a, for CRAM see Sutula et al. 2006, Stein et al. 2009). But, we also conducted our own statistical tests to validate the method, and adjusted the use of metrics before generating a final scorecard (see below). The IEI was generated from the Ecological Observations database after all metrics had been scored in the field. The IEI is presented in a summary scorecard that shows the scoring of all metrics, Major Ecological Factors (MEFs) and primary Rank Factors. Most metrics within each MEF typically received a weight of 1.0. Some were weighted as 0.5 if they were known to be partially redundant with other metrics (e.g., native species cover and invasive species cover metrics) or not as responsive as other metrics. Each MEF received a weight of 1, except for Soils, which received 0.5. Finally, the overall Rank Factors were weighted as follows: Landscape Context – 0.25, Size – 0.15, and Condition – 0.60, based on best professional judgment of how these factors contribute to the overall condition, resistance and resilience of a wetland, as reported in Faber-Langendoen et al. (2011a).

We exported all of the field values from the Databases into formats suitable for analyses. We evaluated the data at multiple indicator levels: metrics, MEFs, and Rank Factors. Each site was a row in a spreadsheet, and each column contained classification and other attribute information as well as the various levels of indicators (over index scores, rank factor and major ecological factor scores, as well as individual metric or indicator scores, and human stressor index scores).

Statistical Screening of Metrics, Attributes and Index of Ecological Integrity

A central question is whether our assessment methods properly grade the wetlands from Excellent to Poor across different wetland types. To evaluate the IEI, we conducted both statistical tests and heuristic evaluations based on previous natural heritage ranks.

Typically we relied on overall rank or categorical values of individual metrics, and used non-parametric analyses. This is because our field observations typically involved assigning a rating of A, B, C, and sometimes D or E to a metric, rather than a numeric score. In addition, we assigned the overall IEI ratings using two variants: A, B, C, CD, and D (5 ratings), or to treat the CD as a subset of C (i.e., 4 ranks = A, B, C (including CD), and D). But for some analyses we chose to treat CD as part of D (4 ranks = A, B, C+ and CD/D), given the paucity of fully D-ranked sites. Could this be said more clearly as "using two variants: A, B, C, CD, and D (5 ratings). Although sometimes CD ranks were lumped with C ranks or D ranks due to the paucity of sites ranked CD or D."

Human Stressor Index

We developed a Human Stressor Index (HSI), following Rocchio (2007), based on a rollup of the individual stressors impact scores reported by field crews for buffer, soils and hydrology (see Table C3). The HSI is = Soils Rating + Hydrology Rating + Buffer Rating / 3. This stressor-based index is primarily independent of the ecological integrity metrics, though some integrity metrics partly consider stressors. To ensure complete independence in the analysis, we also compared the HSI against Vegetation metrics alone.

TABLE C3. Rollup procedure for creating an overall stressor rating for Buffer, Soils, and Hydrology, which together are used to create a Human Stressor Index (see text).

Impact Values of Stressor Categories	OVERALL STRESSOR RATING (points)
1 or more Very High Stressors, OR 2 or more High, OR 1 High + 2 or more Medium	Very High (1)
1 High Stressor, OR 3 or more Medium, *OR* 2 Medium + 2 or more Low, *OR* 1 Medium + 3 or more Low	High (2)
1 Medium Stressor + 5 or more Low Or 8 or more Low	Medium (3)
1 Medium + 1–4 Low 1–7 Low Stressors	Low (4)
0 Stressors	Absent (5)

Screening Metrics

We screened and scored the L2 and L3 metrics using comparable methods for other Indices of Biotic Integrity (IBIs) (Barbour et al. 1996, Blocksom et al. 2002, Klemm et al. 2003, Jacobs et al. 2010). We examined discriminatory power or responsiveness, and redundancy.

Discriminatory power is the ability of a metric to distinguish high stress from low stress sites, based on the Human Stressor Index (HSI), which integrates stressor scores for hydrology, soils, and buffer into an overall score, which is then converted to a stressor rating (High, Medium, Low, Absent). We compared how well various components of the EIA were able to distinguish these sites.

We evaluated metrics by examining their distributions using box-and-whisker plots and one-way analysis of variance (ANOVA) (appropriate to categorical data for L2) to see if any of the metrics had significantly different mean values among the four levels of stressors. We used the **Kruskal-Wallis rank sum test,** which is the non-parametric analogue of a one-way ANOVA (when there is one nominal variable and one measurement variable and the measurement variable does not meet the normality assumption of an ANOVA). The Kruskal-Wallis test does not make assumptions about normality. Like most non-parametric tests, it is performed on ranked data, so the measurement observations are converted to their ranks in the overall data set: the smallest value gets a rank of 1, the next smallest gets a rank of 2, and so on. We applied the **Pairwise Wilcoxon Rank Sum Tests** to calculate pairwise comparisons between group levels with corrections for multiple testing, making adjustments to p values when testing multiple comparisons. Finally, comparisons were scanned using **Notched Boxplots.** If the two boxes' notches do not overlap, this is "strong evidence" that their medians differ (Chambers et al, 1983, p. 62). Variables with non-significant F-values from the ANOVA were considered non-responsive and candidates for removal.

Redundancy was evaluated with the aim of minimizing metrics that were redundant. A key concern for Level 2 evaluations is to keep the assessment time as efficient as possible. Metrics that duplicate other metrics are candidates for elimination. We evaluated metrics and Major Ecological

Factors (by calculating a Spearman rank correlation coefficient matrix among all metrics and MEFs, and reviewing scatterplots (in part to be aware of non-linear patterns). Metrics correlated with an r [s] > 0.8 were considered largely redundant (Jacobs et al. 2010), metrics correlated with an r [s] 0.6–0.79 were considered partially redundant.

When redundant attributes are identified, the one with the strongest correlation to human disturbance and most effective discriminatory power were considered most valuable to retain. If redundant attributes (e.g. woody regeneration and % non-native species) provided unique ecological information (level of abundance of woody saplings and seedlings vs. change in abundance of non-native species) they were retained.

Level 2 and Level 3 metrics (Coefficient of Conservatism)

Mean CC

We used the simple form of the coefficient of conservatism (CC): presence/absence, based on testing by Rocchio (2007). We assessed the relationships between mean CC scores calculated from Level 3 plot data (88 sites) and that of individual L2 metrics, overall vegetation scores, condition, and overall integrity for those some 88 sites. We also evaluated whether the mean CC should be used as a supplement to the more rapid L2 metrics, as one way to add a robust metric for rapid assessments. That is, we assessed whether it was redundant with L2 metrics, and whether it added value as a more quantitative measure.

We did not pursue testing or development of a Vegetation Index of Biotic Integrity (VIBI). VIBIs are available for forest, shrub and herb wetlands in Ohio in some of the same ecoregions and macrogroup as sampled here (Mack 2004, 2007a) and initial testing could start by using those VIBIs.

Vegetation Structure – Floodplain and Swamp Forests

A number of Level 3 metrics are available to us from forested sites, but we are not able to evaluate them at this time, nor are they generally applicable, so they are not scored here. Typical Level 3 metrics could include:

- Structural stage (assessment of the old growth status of forested wetlands) (tree stems > 30 cm dbh, > 50 cm dbh)
- Overstory tree basal area
- Overstory tree density
- Sapling density
- Volume of coarse woody debris

Testing and Applying the Revised EIA Method

Based on revisions to the EIA method from statistical tests, we then assessed how responsive the revised attributes and indices were to the Human Stressor Index. We summarized the IEI scores for our 277 sites.

C.3. Results

We obtained data from 277 sites, of which all contained L2 data and 88 contained L3 data. We focus on the overall patterns of condition and ecological integrity ratings, across all macrogroups.

Statistical Screening – Redundancy

We screened 23 primary metrics, including those that were components of a metric index (e.g., the Connectivity metric index contains two primary (or sub-) metrics—core connectivity and supporting connectivity). We also screened the metric indices (connectivity, land use, buffer), for a total of 26 metrics.

Landscape Context Metrics

The submetrics within both the Land Use Index and the Connectivity Index (core versus supporting) were moderately correlated (and thus only partially redundant) (r = 0.70 for Connectivity, r = 0.71 for Land Use), supporting the use of these two scales. However, the indices themselves, were strongly correlated (and thus considered redundant) (r = 0.89). This is not too surprising given that both metrics rely on similar information from the imager—Connectivity assesses the percent natural ecosystems in the surrounding landscape, and Land Use Index further assesses the intensity of land uses within the cultural part of the landscape. Connectivity is straightforward and simple to measure and would appear to be the preferred metric, at least in this region of the U.S. However, the range of land uses encountered in our study (typically rural to wild landscapes) may not have been widely representative, and we suggest that land use be retained until testing across a greater variety of landscapes is completed.

The Buffer Index was only moderately correlated (and thus only partially redundant) with either Connectivity (r = 0.69) or Land Use (r = 0.68). Given the distinctive evaluation needs and characteristics of the buffer (and its extent, width and condition being assessed), we suggest that buffer should be treated as its own MEF, separate from other Landscape Features.

Size

Relative Size (percentage reduction in wetland size from draining, filling etc. compared to its natural size) and Absolute size were minimally correlated (r = 0.57), suggesting they were only minimally redundant, and provide independent information on Size. Absolute Size was typically not completed by field crews (147 sites not rated), as further research would have been needed to determine optimal size rating scales for wetlands in these ecoregions.

Condition

Within Condition, the MEFs are only minimally correlated (r = 0.31–0.50), indicating that they are minimally redundant, and thus worth assessing separately (Table C4). Of the 3 MEFs, Soils showed the lowest correlation (r =0.50) with Condition, and Vegetation the highest (r = 0.90).

TABLE C4. Correlation (spearman's rank) between the three Major Ecological Factors of Condition and Mean Coefficient of Conservatism (CC).

	Mean CC	Condition	Vegetation	Hydrology	Soils
Mean CC	1.00				
Condition	0.55	1.00			
Vegetation	0.61	0.90	1.00		
Hydrology	0.31	0.78	0.50	1.00	
Soils	0.25	0.50	0.31	0.36	1.00

Composition was the single metric strongly correlated with the overall Vegetation rating (r = 0.88) and with Condition (r = 0.83), suggesting that field evaluation of this metric provided a strong indication of the overall Condition of the site. Mean CC was moderately strongly correlated with Vegetation (r = 0.61), but showed little to no correlation with Hydrology or Soils.

VEGETATION

Within Vegetation, only one pair of metrics was strongly correlated (and thus redundant); percent natives and percent invasive exotics (r = 0.82) (Table C5). This is not surprising, given they are essentially two sides of the same coin, the one from the perspective on condition (percent natives), the other that of the stressor (invasives). Overall Composition was moderately correlated with Structure, Natives, and Invasives (r = 0.61–0.64). Organic Matter had little correlation with any other metric within Condition (r = 0.04–0.37). Increasers also had low correlations (r =-0.10–0.46). Thus for future assessments, we suggest that Organic Matter and Increasers could be dropped (see below for summary). Although either the Natives or Invasives metrics could be dropped, given their correlation, they do provide important perspectives on the condition of vegetation (e.g., some exotics present may not be invasive).

TABLE C5. Correlations among vegetation metrics. Correlations above 0.60 are highlighted

	Structure	**Organic**	**Increaser**	**Native**	**Invasives**	**Regeneration**	**Compo-sition**	**Mean CC**
Structure	1.00							
Organic	0.43	1.00						
Increaser	0.32	0.04	1.00					
Native	0.43	0.09	0.33	1.00				
Invasive	0.35	0.07	0.30	0.82	1.00			
Regeneration	0.40	0.37	0.16	0.31	0.17	1.00		
Composition	0.64	0.32	0.46	0.63	0.61	0.47	1.00	
Mean CC	0.40	0.26	0.38	0.38	0.42	0.53	0.50	1.0

Mean CC showed minimal correlations with other metrics, the highest being with the Composition (r = 0.50) and regeneration (r = 0.53) metrics. As noted above, it did show a moderate correlation with overall Vegetation scores, suggesting that it more strongly reflects the combination of vegetation metrics, rather than any one metric.

HYDROLOGY

The three metrics were only minimally correlated (r = 0.25–0.46) (and thus minimally redundant). The three metrics were minimally correlated with overall Condition (r = 0.56–0.67), though collectively the correlation between Hydrology and overall Condition was fairly strong (r = 0.78) (Table C4 above).

SOILS

The three metrics were uncorrelated with each other (r = 0.12–0.15), and only minimally correlated with overall Condition (r = 0.25–45), with patch diversity having the strongest correlation.

Statistical Screening – Discriminatory Power

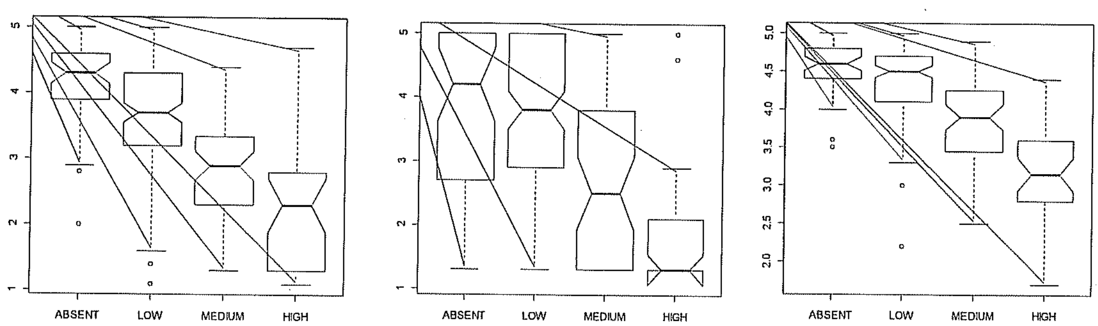
Screening Rank Factors and Attributes

We compared how well various components of the EIA were able to discriminate between sites with different levels of stressors, based on the Human Stressor Index (HSI), from High (including both High and Very High) to Absent (Table 6). Landscape Context, Size, and Condition discriminated among each of the stressor levels Within Condition, Soils showed only weak discrimination, separating highly stressed sites from others. Hydrology and Vegetation were very effective (Table C6, Figs. C6 and C7).

TABLE C6. F-values for one-way Kruskal-Wallis rank sum test on the four stressor levels (High, Medium, Low, Absent) for MEFs and metrics. Significance of F values are shown, and all pairwise comparisons are significant at $p < 0.001$ (<<<), $p<0.01$ (<<) or $p <0.05$ (<). Non adjacent means are equal if underlined. Discriminatory Power ranges from High (all Stressor Levels distinguished) to Poor (no Stressor levels distinguished).

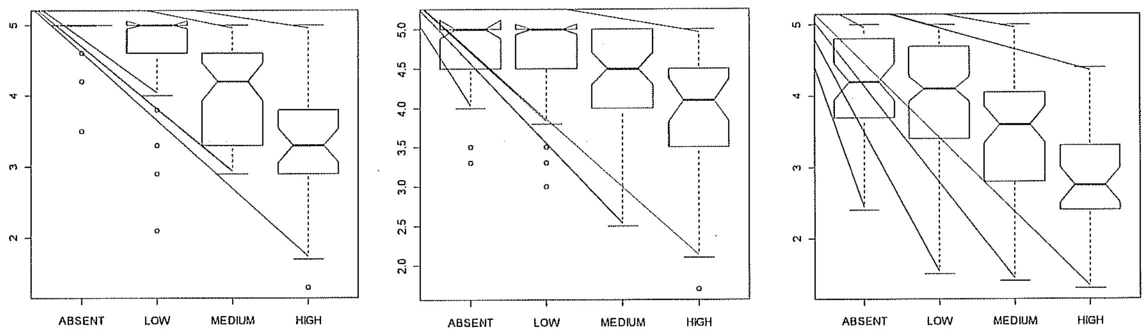
	F-value	**Pairwise comparison**	**Discriminatory Power (H,M,L, P)**
EIA Score	124.7***	A>>>L>>>M>>>H	H
EIA Rank	128.7***	A>>>L>>>M>>>H	H
Landscape	89.5***	A>>L>>>M>>H	H
Size	59.4***	A>>>L>>>M>>>H	H
Condition	109.6***	A>>L>>>M>>>H	H
Hydrology	129.2***	A>>L>>>M>>>H	H
Soil	33.6***	A=L=M>H	L
Vegetation	59.6***	A=L>>>M>>H	M
Mean CC	10.8*	A=L=M=H	P

FIGURE C6. Notched Boxplot of Rank Factor Ratings (x axis) in relation to Human Stressor Index (y axis).



Landscape Context Size Condition

FIGURE C7. Notched Boxplot of MEF Ratings (x axis) within Condition in relation to Human Stressor Index (y-axis).



Hydrology   Soils Vegetation

Screening Metrics

We assessed each metric in turn, to determine its discriminatory power with respect to response to stressors. We used the following ratings:

High (H) = F value p <0.001, and at least 4 stressor levels distinguished.

Moderate (M) = F value p <0.001 and at least 3 stressor levels distinguished

Low (L) = F value p < 0.001 and at least 2 stressor levels distinguished.

Poor (P) = F value p < 0.05 and 1–2 stressor levels distinguished.

All metrics, MEFs and Rank Factors had statistically significant contributions in discriminating among sites categorized by stress. Our concern here is to ascertain those that make the most ecologically important contribution to discriminating among these sites, recognizing that responses to abiotic stressors are not necessarily the entire picture.

Landscape Context Metrics

All three landscape context metrics had moderate to high discriminatory power, leading to the overall high discriminatory power of the rank factor (Table C7). Because all metrics are based on aerial photo interpretation, all metrics were scored for all sites.

TABLE C7. Discriminatory power of Landscape Context Metrics in response to the Human Stressor Index. F-values for one-way Kruskal-Wallis rank sum test on the four stressor levels of the Human Stressor Index (High, Medium, Low, Absent) for metrics. Significance of F values are shown, and all pairwise comparisons are significant at $p < 0.001$ (<<<), $p<0.01$ (<<) or $p <0.05$ (<). Non adjacent means are significantly different unless underlined.

	F-value	Pairwise comparison	Discriminatory Power (H,M,L, P)
LANDSCAPE	89.5***	A>>L>>>M>>H	H
Connectivity	62.0***	A>>L>>>M>>H	H
Land Use Index	74.0***	A>>L>>>M>>H	H
Buffer Index	80.1***	A>>>L>>>M=H	M

Size Metrics

The two size metrics showed moderate discriminatory power, and together provided an overall high discriminatory power for the rank factor and attribute (Table C8). But field crews also had difficulty scoring absolute size; 147 sites were not given ratings. 35 sites were not given Relative Size ratings.

TABLE C8. Discriminatory power of Size Metrics based on Human Stressor Index. See Table C7 for details.

	F-value	Pairwise comparison	Discriminatory Power (H,M,L, P)
SIZE	59.4***	A>>>L>>>M>>>H	H
Absolute Size	33.77***	A=L>>M>H	M
Relative Size	54.4***	A=L>>M>>H	M

Condition Metrics

Condition metrics were organized by Major Ecological Factors of Vegetation, Hydrology, and Soils. Vegetation overall had moderate discriminatory power to the abiotic stressor index; only the Composition metric had a moderate discriminatory response, the rest had low (Table C9). Regeneration was left blank or assigned a Non Applicable rating at 191 sites. *Hydrology* overall had high discriminatory power, with all metrics having moderate discriminatory power. All metrics were scored at all sites. *Soils* overall had low discriminatory power; patch diversity had a moderate rating, the others poor, including Water quality, which was left blank or assigned a Not Applicable rating at 56 sites.

Given that the stressor index is based on abiotic stressors, it is perhaps not surprising that hydrology had the best discriminatory power. Vegetation metrics appeared to be only partly responding to the assessed abiotic stressors. However, soil metrics appeared to be of little value in assessing response to stressors.

TABLE C9. Evaluation of Discriminatory Power of Condition Metrics based on Human Stressor Index. See Table C7 for details.

	F-value	Pairwise comparison	Discriminatory Power (H,M,L, P)
CONDITION	109.6***	A>>L>>>M>>>H	H
Vegetation	59.6***	A=L>>>M>>H	M
Structure	37.1***	A=L>>M=H	L
Regeneration	20.3***	A=L>>M=H	L
Organic Matter	26.6***	A=L>M=H	L
Natives	25.2***	A=L>M=H	L
Invasives	23.3***	A=L>M=H	L
Increasers	15.0**	A=L=M>H	L
Composition	59.9***	A=L>>>M>>>H	M
Hydrology	129.2***	A>>L>>>M>>>H	H
Water Source	59.5***	A>>L>>>M=H	M
Hydro. Connectivity	83.1***	A=L>>>M>>H	M
Hydroperiod	98.6***	A=L>>>M>>>H	M
Soil	33.6***	A=L=M>H	L
Patch Diversity	30.0***	A=L>M>M	M
Soil Surface	10.2*	A=L=M=H	P
Water Quality	8.6*	A=L=M=H	P
Mean CC	10.8*	A=L=M=H	P

Final Selection of Metrics

Vegetation metrics and MEFs were reviewed once more to ensure correlations were not based on outliers and that each was ecologically meaningful. We compared pairs of metrics that were redundant, and those that had the strongest correlation to human disturbance and the most effective discriminatory power, based on the 2nd part of our screening, were considered most valuable to retain. Only two pairs of metrics were flagged as redundant. The first pair, connectivity and land use, were both highly correlated. Both also showed moderate discriminatory power. The 2nd pair, native and invasives were both highly correlated, but only of lower discriminatory power with respect to abiotic stressors. For now we reduced the weight of each metric to 0.5, and in the future suggest they could be combined into a Native-Invasives Species Index, as there is value in knowing both the total native species cover and the cover of invasives.

In terms of discriminatory power, we proceeded cautiously, because discrimination was based on abiotic stressor gradient, and not all integrity changes are due to abiotic stressors. If some redundant or low discriminating metrics (e.g. woody Regeneration and Invasive Species) provided unique ecological information (level of abundance of woody saplings and seedlings vs. change in abundance of non-native species) they were retained.

Based on this review, we simplified our set of metrics for wetland EIAs, reducing the total number from 18 to 15, with 1 optional metric, as shown in Table C10. For Landscape Context, we suggest that Land the Use Index is optional, but before we can recommend dropping it, we would like to test it across more land use types in different regions of the country. Size remains unchanged. For Vegetation, we removed the Organic Matter and Increasers metrics. Hydrology remains unchanged. For Soils, we removed the Water Quality metric.

TABLE C10. Revised set of metrics for wetland Ecological Integrity Assessments. Rd = Redundancy. DP = Discriminatory Power.

RANK FACTORS	MAJOR ECOLOGICAL FACTORS	METRICS (original)	Rd	DP	METRICS (revised)
LANDSCAPE CONTEXT	LANDSCAPE	Connectivity	Y	H	Connectivity
		Land Use Index	Y	H	Land Use Index ***(optional)***
	BUFFER	Buffer Index		M	Buffer Index
SIZE	SIZE	Relative Patch Size (ha)		M	Relative Patch Size (ha)
		Absolute Patch Size (ha)		M	Absolute Patch Size (ha)
CONDITION	VEGETATION	Vegetation Structure		L	Vegetation Structure
		Organic Matter		L	***Removed***
		Regeneration (woody)		L	Regeneration (woody)
		Native Plant Species Cover	Y	L	Native Plant Species Cover [weight 0.5]
		Invasive Plant Spp. – Cover	Y	L	Invasive Plant Spp. – Cover [weight 0.5]
		Increasers – Cover		L	***Removed***
		Vegetation Composition		M	Vegetation Composition
	HYDROLOGY	Water Source		M	Water Source
		Hydroperiod		M	Hydroperiod
		Hydrologic Connectivity		M	Hydrologic Connectivity
	SOIL	Physical Patch Types		M	Physical Patch Types
		Soil Disturbance		P	Soil Disturbance
		Water Quality		P	***Removed***

Applying the Final Model

Comparing Integrity Scores to the Stressor Index

We examined the relationship between the ecological integrity scores and the stressors acting on the site, both on site and observed in the buffer using both the revised model, using the Human Stressor Index to examine relationships.

We found a moderately strong correlation between the HSI and the IEI (r = 0.66), suggesting that our measures of ecological integrity do moderately well at showing responses to human stressors (Table

C11). The condition rank factor showed the greatest response, and within Condition, the Hydrology attribute showed the strongest response. This may reflect how stressors to hydrology are the most readily observable (e.g., ditches, dikes), as are their effects on hydrologic condition.

Neither Mean CC nor EORANK showed a particularly strong relation to the HSI,

All individual metrics showed minimal correlations with the HSI (all $r < 0.60$), suggesting that multiple metrics are needed to assess the combinations of stressors acting on a site.

TABLE C11. Correlation of ecological integrity rank factors and major ecological factors to the Human Stressor Index based on the final set of metrics for the EIA model.

RANK FACTOR **Major Ecological Factor**	**HSI Rating**
LANDSCAPE	0.57
Landscape	0.50
Buffer Index	0.54
SIZE	0.46
CONDITION	0.61
Vegetation	0.45
Hydrology	0.67
Soil	0.32
Index of Ecological Integrity	0.66
Mean_C	0.30
EORANK	0.34

Comparing Landscape Context and On-Site Condition

We assessed the degree to which on-site Condition ratings were correlated to surrounding L1 Buffer and Landscape metrics. The overall Landscape Context Rank Factor had a modest correlation ($r = 0.47$) with on-site Condition, and among the individual metrics, the Buffer Index was the strongest ($r=0.48$) (Table C12). Within overall Condition, Hydrologic Condition was most strongly correlated with Landscape Context and its metrics ($r = 0.52$–0.54). In addition, the Human Stressor Index had a modest correlation with Landscape Context metrics (0.52–0.57).

TABLE C12. Landscape Context metrics versus Condition and Stressor Index ratings. The Landscape Context metrics are based on Level 1 metrics; the Condition and Stressor Index on Level 2 metrics.

	LANDSCAPE CONTEXT	Metrics		
		Connectivity	Land Use	Buffer Index
CONDITION	0.47	0.37	0.41	0.48
Vegetation	0.34	0.26	0.28	0.36
Hydrology	0.54	0.46	0.48	0.52
Soils	0.18	0.13	0.15	0.21
Mean_CC	0.40	0.39	0.34	0.37
HSI	0.57	0.47	0.52	0.54

Level 3 (Mean CC) and Level 2 Metrics

Our evaluation of the responsiveness of the mean CC values (from 88 sites) showed that it did not provide a readily interpretable part of on-site condition or overall Ecological Integrity. It added little to what we had learned from the Level 2 metrics (scored across all 277 sites).

Level 1 Predictions of Level 2 Ratings

We tested how well Level 1 ratings predict Level 2 ratings at individual sites. We compared two approaches that used the Landscape Context metrics around a site to predict on-site condition. The first approach is based on the general Landscape Condition Model of Comer and Hak (2009) (Fig. 8a), as summarized in Appendix A; the 2nd is based on our Landscape Context metrics of the EIA model used here. Both models do a reasonable job of separating out differences in Condition, but the general Landscape Condition Model is less successful than the Landscape Context Ratings, including being able to recognize only 3 levels of condition, rather than the 4 levels of the latter (Figs C8a, C8b). However, the models are more successful in predicting overall IEI scores (not shown), because landscape context and size, as well as on-site condition are relevant to an IEI, and level 1 models do well at assessing these aspects of integrity. Although not attempted here, it may also be possible to recalibrate the Level 1 metrics used for Level 1 assessments based on these Level 2 scores in order to improve the prediction of Level 2 Condition from Level 1 metrics.

FIGURE 8. Correlation between two kinds of Level 1 EIA (remote sensing) models and Level 2 (field-based) assessments of On-site Condition (with scale of A (5.0) to D (D= 1.25). nd the Level 1 EIA Method has a combination of stressor and integrity based metrics.

FIGURE 8a. The Landscape Condition Model rating (based on stressor-based metrics) for the combined area of wetland and core landscape (1 km buffer around wetland). Kruskal-Wallis F = 32.6, p < 0.001. VL= L< M< H =VH.

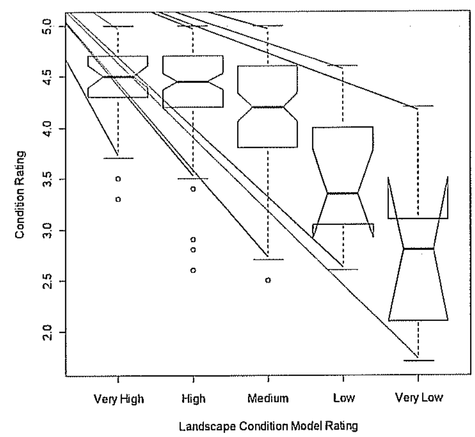
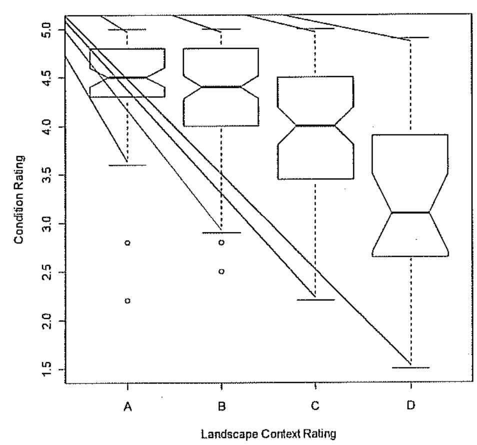


FIGURE 8b. Level 1 EIA Landscape Context ratings (based on three landscape context metrics of Table C10). Kruskal-Wallis F = 52.0, p < 0.001. A>B>> C>>D.



Comparison with Natural Heritage Ranks

Our presumption in designing the study was that previous evaluations of wetland condition by network ecologists were somewhat inconsistent, but in line with the ecological integrity approach. We would like to determine more carefully just how the two are related. But it would be hard to use the provided EORANKs. This is because previous network evaluations may have occurred anytime in the last 2 to 20 years, so the field conditions of the wetlands may have changed from the previous visit. Instead, the MNFI staff ecologists decided to revise their EORANKS based on the visits over the last two years using their standard methods. There is some potential circularity here because they participated on the field teams completing the EIAs. But they had no knowledge of the preliminary or final IEI ratings assigned here. Thus we felt the comparison would still be helpful. We compared the findings of our IEI with those of the original ranks assigned by natural heritage programs using Spearman's rank correlation coefficients ($r_{[s]}$). Spearman's rank is used because the ranks are ordinal data.

Michigan re-ranked 125 sites based on the 2009–2010 data. We first note that, for Michigan, the correlation between the natural heritage ranks pre-2009 with those of 2009–2010 was $r = 0.48$ ($p < 001$). Despite the statistically significant correlation, it was not a high correlation. Thus the new MNFI ranks changed considerably from the previous ratings.

Overall correlation between IEI numerical score and the new MNFI EORANK was $r = 0.62$, whereas correlations with the IEI categorical rating and MNFI EORANK was 0.53. Michigan Condition rating correlated most strongly with the EIA Vegetation rating ($r = 0.70$), and, more surprising, was correlated $r = 0.70$ with the Composition metric rating alone! By adding in Hydrology and Soils into the IEI rating, the correlation between Michigan Condition and EIA Condition drops to a moderate correlation ($r = 0.44$). Michigan Condition rating and mean CC had no significant correlation ($r = 0.12$), whereas EIA Vegetation Rating and mean CC had a moderate correlation ($r = 0.49$).

For the HSI rating, the only Michigan factor showing any correlation with the stressor index was Michigan Landscape Context ($r = 0.49$). The overall MNFI EORANK had low correlation ($r = 0.28$). By comparison, IEI Rating had a moderate correlation ($r = 0.47$). So, our EIA method assesses a broader range of ecological attributes, and these show a response to stressor impacts.

Application of IEI Scores to Wetland Sites

We used our final form of the IEI, along with ratings for each of the Major Attributes to summarize the reference gradient for wetlands in Michigan and Indiana (Table C13). We summarize the ecological characteristics, the ecological integrity ratings, and stressors across the major wetland types in terms of Excellent (A), Good (B), Fair (C), and Poor (D) wetlands.

TABLE C13. Summary of the number of sites by wetland macrogroups and a) Index of Ecological Integrity rating, and b) Condition. Highlighted cells are those with less than 5 replicates.

	Index of Ecological Integrity					
Macrogroups	**A**	**B**	**C**	**CD**	**D**	**Total**
Northern & Central Floodplain Forest	8	19	5	6	1	39
Northern & Central Swamp Forest	5	21	5	2	1	34
North American Boreal Bog & Fen	12	17	0	0	1	30
Great Plains Freshwater Wet Meadow, Wet Prairie & Marsh	6	12	8	7	5	38
Appalachian, Interior Plateau & Prairie Fen	6	32	2	2	2	44
Eastern North American Wet Shrub, Meadow & Marsh	6	31	7	9	1	54
Atlantic & Gulf Coastal Plain Pondshore & Wet Meadow	17	16	3	2	0	38
Total	60	148	30	28	11	277

	Condition				
Macrogroups	**A**	**B**	**C**	**D**	**Total**
Northern & Central Floodplain Forest	17	21	1		39
Northern & Central Swamp Forest	17	14	2		33
Appalachian, Interior Plateau & Prairie Fen	20	21	3		44
North American Boreal Bog & Fen	24	5	1		30
Atlantic & Gulf Coastal Plain Pondshore & Wet Meadow	26	9	3		38
Eastern North American Wet Shrub, Meadow & Marsh	12	36	7		55
Great Plains Freshwater Wet Meadow, Wet Prairie & Marsh	11	20	6	1	38
Total	127	126	23	1	277

C.4. Discussion

The Ecological Integrity Model

Our overall model for wetlands was developed from a large body of wetland condition assessment methods. Thus the choice of metrics and the variants needed for various wetland types (bogs, marshes, floodplain forests) was based on previous testing. Being something of a synthesis of other models, we were concerned that we might have included too many metrics, or that some metrics were not as relevant to assessing integrity as they might be to other aspects of ecosystem structure and function. In addition, we included metrics, such as Vegetation Composition, that are less quantitative, and require greater ecological experience than other metrics, such as Percent cover of Natives. For that reason, we felt it was important to check the model for redundancy and discriminatory power.

By and large, the individual metrics, as well as aggregate major ecological factors,primary rank factors, and the overall index of ecological integrity showed little redundancy and good discriminatory power. Still, of the 18 major metrics we included, we felt we could do with as few as 14 (one to four metrics per six major attributes). Reducing the number of metrics will increase the efficiency of the overall process.

In evaluating the contribution of various ecological attributes to understanding ecological integrity, we found that Vegetation was not a simple proxy for Hydrology or Soils, or does not respond strongly to abiotic stressors (Table C11). Our tests used a Human Stressor Index based on Stressors to the abiotic (Hydrology and Soils) and buffer attributes, but the vegetation metrics responded somewhat weakly to those stressors ($r = 0.45$). Changes in the vegetation may reflect other stressors (e.g. logging history, deer browse) not assessed by the on-site abiotic stressors, or reflect longer term response to stressors that are not currently evident on the site. The relative weight given to the Vegetation attribute may vary depending on the application. From a biodiversity perspective, some may feel that, even if hydrology and soils are not degraded, but vegetation is strongly degraded, then the IEI should be rated as poor. From a functional perspective more weight could be given to hydrology or soils. Here we chose a balanced set of weightings.

Our model, more than some, includes metrics that require experienced wetland ecologists to properly rate them, much as wetland delineation requires experienced evaluators. We believe the use of such metrics to be the strength of the method, allowing us to retain at least a four point rating of integrity (A, B, C, D), as opposed to A/B, C and D.

Our field assessments rely on observable features at the site. But past land use history may play an important role in shaping current condition. This includes activities that have occurred (or not) on a particular site anywhere from a few years ago to 200 years ago. Their influence may be unknown or very uncertain to, and yet they may affect the level of ecological integrity. For example, a site may have altered hydrology, but when combined with decades of very benign land use, they maintain a good condition. There may also be lag effects from land uses. Impacts from crop production (edge effects from pesticides, nitrogen enrichment and runoff) may show up right away in condition, whereas fire suppression might not show up strongly until a tipping point is reached, maybe after decades or even 100 years.

Overall Level 2 Index and Level 3 metrics

By constructing a model that consistently assesses the major attributes of a wetland, we are able to identify those aspects of the wetland that are in better or worse shape. Thus, we are able to show here that few wetlands have soil disturbances, but many have altered hydrology and vegetation.

Integrating these various attributes into an overall rank provides a concise and readily interpretable rating of wetland integrity for wetland managers, conservationists, and the public.

The IEI is based on the perspective of a minimally disturbed reference condition, based on historical integrity of the site, lack of negative human impacts, and a surrounding landscape dominated by natural ecosystems and processes. But a high score for integrity may not necessarily translate into high scores for ecosystem services. That is a separate evaluation. We caution that wetlands may be in excellent condition but not may not be considered high scoring for any given ecosystem service. Thus, ecosystem services should be evaluated based on both the inherent capacity of natural ecosystems as well as potential capacity based on modifications to those systems. For example, floodplain forests with high ecological integrity have a range in capacity for providing flood control services; these forests could also be modified to increase those services, but depending on the modification, this may or may not maintain their level of integrity.

Although our weighting of metrics is general, it has the advantage that it can be applied to any wetland, at least in the temperate and boreal regions. There are advantages to fine-tuning the model, weighting some metrics or attributes higher than others, depending on the wetland. For example, Jacobs et al. (2010) found some justification for weighting buffer more strongly than hydrology for depressional wetlands, and vice-versa for riverine and flats. Nonetheless, applying these rules by wetland type resulted in only minimal gains in the sensitivity of the index, using intensive quantitative metrics (Jacobs et al. 2010, Table 5). Thus for rapid assessments, we feel a general model will suffice, with Condition weighted 60% (Vegetation 24, Hydrology 24, and Soils 12), Size 15%, and Landscape Context 25%. Thus all major attributes are scored roughly in the 12–25% range. This is in keeping with the IEI model, in that these attributes are included precisely because they reflect major attributes of the wetland, and as our data show, are each responsive to stressors.

The Level 3 metric based on Coefficient of Conservatism added a little to our understanding of ecological integrity, particularly in validating our Level 2 Vegetation metric ($r = 0.61$). Among more specific Vegetation metrics, Mean CC showed the highest correlation with Composition ($r = 0.50$) and regeneration ($r = 0.53$) metrics, suggesting some support for the expert evaluation of vegetation. Further analyses of the Level 3 data are needed. Also, though not the purpose of this study I think the following is true: Level 3 assessments serve to enhance field skills of both beginner and experienced level surveyors; by requiring relatively comprehensive plant species identification and cover estimates, those skills are honed. Conducting the advanced work of such plot studies tend to create stronger interpretive skills for the less advanced Level 2 surveys.

Calibrating IEI with Remote Sensing Models

The Level 1 integrity ranks are often used as a means of prioritizing sites for field visits where Level 2 or Level 3 assessments will be completed (e.g., see Fennessy et al. 2007), and ranks based on those assessments would supersede Level 1 ranks. Thus Level 1 assessments can be informative about the overall range in conditions across a population of wetlands in a landscape or region. They can serve as a helpful screening method for identifying the most likely conditions on the ground. Level 1 ratings can also be used as predictors of Level 2 or 3 ratings at individual sites. Tests completed to date, however, show that Level 1 methods do not accurately predict individual site ratings, particularly on-site conditions (Mack 2006, Fennessy et al. 2007). Our tests for wetland sites in Michigan and Indiana bear this out (Fig. 8). However, these tests shows that our methods are more successful in predicting overall IEI scores, because landscape context and size, as well as on-site condition are part of the IEI, and these can be effectively assessed using Level 1 metrics. It may also be possible to recalibrate the metrics used for Level 1 assessments based on these Level 2 scores.

Study Design – Assessing Wetlands Versus Points

Our sampling of the reference gradient (minimally disturbed to degraded) was designed to assess "wetlands," versus a fixed area around the sample point. In the context of ambient monitoring this decision may be less desirable, at least from some aspects. As Fennessy et al. (2007) note, probabilistic surveys that have been undertaken have taken an area-based approach rather than assessing a "wetland." This approach avoids 1) the need for determining an assessment unit boundary (which can become difficult in large contiguous complexes of wetlands), and 2) measuring the area of the assessment unit. It also allows points to fall onto disturbed and undisturbed areas of wetlands and be separately assessed, which avoids having multiple sample points being dropped on the same "wetland" (since the available digital sample frames will probably not correspond to assessment units defined by the assessment unit rules of the sample protocol). But there are also several distinct advantages to using the polygon approach, as noted by Fennessy et al. (2007): the basic "currency" in Clean Water Act Section 401/404 regulation of wetlands is something called a "wetland" and this is understood to be "a definable piece of real estate that can be mapped and walked around." These advantages are also evident in many conservation and resource management contexts and why the natural heritage methodology routinely requires mapping of ecosystem occurrences.

We found that defining an assessment unit boundary was relatively straightforward at the majority of sites, since most wetlands in the watershed were relatively small in size (< 50 ha). We were able to obtain digitized assessment unit boundaries and area estimates in advance, but occasionally crews came to a site wherein natural community type boundaries, natural community quality boundaries, and jurisdictional wetland boundaries were in different locations, and had to draw the assessment unit directly onto the aerial photo and digitize this by hand in ArcView. Our field manual will need to improve the guidance on defining Assessment Areas to deal with these complex situations (Faber-Langendoen 2011). But given the advantages of assessing wetlands based on polygons, rather than points, we encourage future assessments to use them.

A Standard Method for Assessing Wetland Condition

Building on the work of other rapid assessment methods (ORAM, CRAM), we show here that ecological integrity can be effectively assessed using a suite of rapid assessment metrics, structured around a general ecological model. Although some of our metrics require greater expertise than others, all attributes have at least two metrics that can be evaluated in a relatively straightforward manner, allowing for wide applicability. Many of these metrics are comparable to the draft metrics being tested as part of the USA-RAM (Table A8), suggesting that a standardized rapid assessment method for wetlands can be achieved. This method will have great value for the NatureServe network,as well as an improved method for assessing wetland condition within and among states and provinces.

References

Andreasen, J.K., R.V. O'Neill, R. Noss, and N. C. Slosser. 2001. Considerations for the development of a terrestrial index of ecological integrity. Ecological Indicators 1: 21–35.

Bailey, R.G. 1996. Ecosystem Geography. Springer-Verlag, NY.

Barbour, M.T., J. Gerritsen, G.E. Griffith, R. Frydenborg, E. McCarron, J.S. White, and M.L. Bastian. 1996. A framework for biological criteria for Florida streams using benthic macroinvertebrates. Journal of North American Benthological Society 15(2): 185–211.

Beyer, H. L. 2004. Hawth's Analysis Tools for ArcGIS. Available at *www.spatialecology.com/htools*

Blocksum, K.A. 2003. A performance comparison of metric scoring methods for a multi-metric index for mid-Atlantic Highlands streams. Environmental Management 31: 670–682.

Blocksom, K. A., J. P. Kurtenbach, D. J. Klemm, F. A. Fulk, and S. M. Cormier. 2002. Development and Evaluation of the Lake Macroinvertebrate Integrity Index (LMII) for New Jersey Lakes and Reservoirs. Environmental Monitoring and Assessment 77:311–333.

Brewer, J.S. and T. Menzel. 2009. A method for evaluating outcomes of restoration when no reference sites exist. Restoration Ecology 17:4–11.

Brinson, M. M. 1993. A Hydrogeomorphic Approach to Wetland Functional Assessment. Technical Report WRP-DE-4. Waterways Experiment Station, U.S. Army Corps of Engineers, Vicksburg, MS.

Brooks, R.P., D.H. Wardrop, and J.A. Bishop. 2004. Assessing wetland condition on a watershed basis in the Mid-Atlantic region using synoptic land-cover maps. Environmental Monitoring and Assessment 94:9–22.

Brown,N., L. Master, D. Faber-Langendoen, P. Comer, K. Maybury, M. Robles, J. Nichols and T.B. Wigley. 2004. Managing elements of biodiversity in sustainable forestry programs: Status and utility of NatureServe's information resources to forest managers. Technical Bulletin No. 885. Research Triangle Park, NC: National Council for Air and Stream Improvement, Inc.

Collins, J.N., E.D. Stein, M. Sutula, R. Clark, A.E. Fetscher, L. Grenier, C. Grosso, and A. Wiskind. 2006. California Rapid Assessment Method (CRAM) for Wetlands and Riparian Areas. Version 4.2.3. 136 pp.

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Menard, M. Pyne, M. Reid, K. Schulz, K. Snow, and J. Teague. 2003. Ecological Systems of the United States: A Working Classification of U.S. Terrestrial Systems. NatureServe, Arlington, VA.

Comer, P., and K. Schulz. 2007. Standardized Ecological Classification for Meso-Scale Mapping in Southwest United States. *Rangeland Ecology and Management* 60 (3) 324–335.

Comer, P.J. and J. Hak. 2009. NatureServe Landscape Condition Model. Internal documentation for NatureServe Vista decision support software engineering, prepared by NatureServe, Arlington, VA. [draft]

Cowardin, L. M., V. Carter, F. C. Golet, and E. T. LaRoe. 1979. Classification of the wetlands and deepwater habitats of the United States. U.S. Fish and Wildlife Service, Washington, DC.

Danz, N.P., G.J. Neimi, R.R.Regal. et al. 2007. Integrated measures of anthropogenic stress in the U.S. Great Lakes Basin. Environmental Management 39:631–647.

DeKeyser, E.S., D.R. Kirby, and M.J. Ell. 2003. An index of plant community integrity: development of the methodology for assessing prairie wetland plant communities. Ecological Metrics 3: 119–133.

De Leo, G. A., and S. Levin. 1997. The multifaceted aspects of ecosystem integrity. Conservation Ecology [online] **1:3**:Available from http://www.consecol.org/vol1/iss1/art3.

Egan, D. and E.A. Howell. 2005. Introduction. Pp. 1–23 *In* D. Egan and E.A. Howell. The Historical Ecology Handbook, Island Press, WA.

Faber-Langendoen, D., G. Kudray, C. Nordman, L. Sneddon, L. Vance, E. Byers, J. Rocchio, S. Gawler, G. Kittel, S. Menard, P. Comer, E. Muldavin, M. Schafale, T. Foti, C. Josse, J. Christy. 2008. *Ecological Performance Standards for Wetland Mitigation based on Ecological Integrity Assessments.* NatureServe, Arlington, VA. + Appendices.

Faber-Langendoen, D, J. Rocchio, P. Comer, G. Kudray, L. Vance, E. Byers, M. Schafale, C. Nordman, E. Muldavin, G. Kittel, L. Sneddon, M. Pyne and S. Menard. 2008. Overview of Natural Heritage Methodology for Ecological Element Occurrence Ranking based on Ecological Integrity Assessment Methods [Draft for Network Review]. NatureServe, Arlington, VA.

Faber-Langendoen, D., D.L Tart, and R.H.Crawford. 2009a. Contours of the revised U.S. National Vegetation Classification standard. Bulletin of the Ecological Society of America 90:87–93.

Faber-Langendoen, D, R. Lyons, and P. Comer. 2009b. Developing options for establishing reference conditions for wetlands across the lower 48 states. A report to the U.S. Environmental Protection Agency. NatureServe, Arlington, VA.

Fancy, S.G., J.E. Gross and S. L. Carter. 2009. Monitoring the condition of natural resources in US national parks. Environ. Monit. Assess. 151:161–174.

Federal Geographic Data Committee. 2008. Vegetation Classification Standard, version 2 FGDC-STD-005, v2. Washington, DC.

Fennessy, M.S., A.D. Jacobs, and M.E. Kentula. 2007. An evaluation of rapid methods for assessing the ecological condition of wetlands. Wetland 27:543–560.

Fennessy, M. S., J. J. Mack, E. Deimeke, M. T. Sullivan, J. Bishop, M. Cohen, M. Micacchion and M. Knapp. 2007. Assessment of wetlands in the Cuyahoga River watershed of northeast Ohio. Ohio EPA

Technical Report WET/2007-4. Ohio Environmental Protection Agency, Division of Surface Water, Wetland Ecology Group, Columbus, OH.

Forman, R.T.T. 1995. Land Mosaics: the ecology of landscapes and regions. Cambridge University Press.

Harwell, M.A., V. Myers, T. Young, A. Bartuska, N. Gassman, J. H.Gentile, C. C. Harwell, S. Appelbaum, J. Barko, B. Causey, C. Johnson, A. McLean, R. Smola, P. Templet, and S. Tosini. 1999. A framework for an ecosystem integrity report card. BioScience 49: 543–556.

Heinz Center. 2002. The State of the Nation's Ecosystems: Measuring the Lands, Waters and Living Resources of the United States. The Heinz Center for Science, Economics and the Environment. Cambridge University Press, New York. 270 pp.

Herman, K.D., L.A. Masters, M.R. Penskar, A.A. Reznicek, G.S. Wilhelm, W.W. Brodowicz, and K.P. Gardiner. 2001. Floristic quality assessment with wetland categories and computer application programs for the state of Michigan. Michigan Dept. of Natural Resources, Wildlife Division, Natural Heritage Program, Lansing, MI.

Herrick J.E., J.W. Van Zee, K.M. Havstad, L. Burkett and W. G. Whitford. 2005. Monitoring manual for grassland, shrubland and savanna ecosystems. Vol 1: quick start. Tucson, AZ: University of Arizona Press.

Higgs, E. 2003. Nature by Design; People, Natural Process, and Ecological Restoration. The MIT Press, Cambridge, MA. 341 p.

Higgs, E.S., and R.J. Hobbs. 2010. Wild design: Principles to guide interventions in protected areas. Pp. 234–251 *in* D.N. Cole and L. Yung. Beyond naturalness: Rethinking park and wilderness stewardship in an era of rapid change. Island Press, Washington, DC.

Holling C.S. 1973. Resilience and stability of ecological systems. Annual Review of Ecology and Systematics 4: 1–23.

Homoya, M.A., D.B. Abrell, J.R. Aldrich & T.W. Post. 1985. The natural regions of Indiana. Proceedings of the Indiana Academy of Science 94:245–268

Hruby, T. 2001. Environmental auditing: testing the basic assumption of the hydrogeomorphic approach to assessing wetland functions. Environmental Management Vol. 27, No. 5, pp. 749–761.

Huggins, D.G. and A. Dzialowski. 2005. Identification and Quantification of Reference Conditions Associated with Lotic Ecosystems of the Central Plains and Surrounding Regions: A summary of approaches and factors and regional approach. Report 134 of the Kansas Biological Survey. University of Kansas, Lawrence, KS.

Jacobs, A.D. 2003. Delaware Rapid Assessment Procedure, Version 1.2. Delaware Department of natural Resources and Environmental Control. Dover , DE.

Jacobs, A., E. McLaughlin, and K. Havens. 2010. Mid-Atlantic Tidal Wetland Rapid Assessment Method Version 3.0. June 2010. Delaware Department of Natural Resources and Environmental Control. Dover, DE.

Jacquart, E., M. Homoya, L. Casebere. 2002. Natural communities of Indiana. July 1, 2002 Working Draft. Department of Natural Resources, Division of Nature Preserves. Indianapolis, IN.

Jennings, M.D., D. Faber-Langendoen, O.L. Loucks, R.K. Peet, and D. Roberts. 2009. Standards for associations and alliances of the U.S. National Vegetation Classification. Ecological Monographs 79: 173–199.

Kapos, V, I Lysenko, and R Lesslie. 2002. Assessing forest integrity and naturalness in relation to biodiversity. Forest Resources Assessment Programme, Working Paper 54, Rome 2002. Forestry Department Food and Agriculture Organization of the United Nations. FRA 2000.

Karr, J.R. and E.W. Chu. 1999. Restoring life in running waters: better biological monitoring. Washington (DC). Island Press. 206 pp.

Klimas, Charles V., Elizabeth O. Murray, Henry Langston, Theo Witsell, Thomas Foti, Rob Holbrook. 2006. A Regional Guidebook for Conducting Functional Assessments of Wetland and Riparian Forests in the Ouachita Mountains and Crowley's Ridge Regions of Arkansas. ERDC/EL TR-06-14 U.S. Army Engineer Research and Development Center. Vicksburg, MS.

Kost, M.A., D.A. Albert, J.G. Cohen, B.S. Slaughter, R.K, Schillo, C.R. Weber, and K.A. Chapman. 2007. Natural Communities of Michigan: Classification and Description. Michigan Natural Features Inventory, Report Number 2007-21, Lansing, MI. 314 pp.

Landres, PB, P Morgan and FJ Swanson. 1999. Overview of the use of natural variability concepts in managing ecological systems. Ecological Applications 9(4): 1179–1188.

Lapp, H., Morris, R. A., Catapano, T., Hobern, D. and Morrison, N. (2011), Organizing our knowledge of biodiversity. Bulletin of the American Society for Information Science and Technology, 37: 38–42. doi: 10.1002/bult.2011.1720370411

Lindenmayer, D.B., and J.F. Franklin. 2002. Conserving forest biodiversity: A comprehensive multiscaled approach. Island Press, Washington, DC. 351 p.

Mack, J.J., 2001. Ohio rapid assessment method for wetlands v. 5.0, user's Manual and scoring forms. Ohio EPA Technical Report WET/2001-1. Ohio Environmental Protection Agency, Division of Surface Water, Wetland Ecology Group, Columbus, OH.

Mack, J. J. 2004. Integrated wetland assessment program, part 4: Vegetation Index of Biotic Integrity (VIBI) and Tiered Aquatic Life Uses (TALUs) for Ohio wetlands. Ohio EPA Technical Report WET/2004-4, Ohio Environmental Protection Agency, Wetland Ecology Group, Division of Surface Water, Columbus, OH.

Mack, J.J. 2006. Landscape as a predictor of wetland condition: an evaluation of the Landscape Development Index (LDI) with a large reference wetland dataset from Ohio. Environmental Monitoring and Assessment 120: 221–241.

Mack, J.J. 2007a. Developing a wetland IBI with statewide application after multiple testing iterations. Ecological Indicators 7: 864–881.

Mack, J.J. 2007b. Integrated wetland assessment program, part 9: Field manual for the Vegetation Index of Biotic Integrity for wetlands. Ohio EPA Technical Report WET/2007-6, Ohio Environmental Protection Agency, Wetland Ecology Group, Division of Surface Water, Columbus, OH.

Mack, J. J. and M.E. Kentula. 2010. Metric Similarity in Vegetation-Based Wetland Assessment Methods. EPA/600/R-10/140. U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.

McRae, B.H., B.G. Dickson, T.H. Keitt, and V.B. Shah. 2008. Using circuit theory to model connectivity in ecology and conservation. Ecology 10: 2712–2724.

McCune, B., and Grace, J.B. 2002. Analysis of ecological communities. MJM Software Design. Gleneden Beach, OR.

McCune, B, and Mefford, M.J. 2006. PC-ORD Multivariate Analysis of Ecological Data, Version 5.0,MjM Software Design, Gleneden Beach, OR.

MDEQ (Michigan Department of Environmental Quality. 2008. Michigan Rapid Assessment Methods for Wetlands (MiRAM), Version 2.0. MDEQ, Land and Water Management Division, Wetlands, Lakes and Streams Unit, Lansing, MI.

Millar CI, Stephenson NL, Stephens SL (2007) Climate change and forests of the future: managing in the face of uncertainty. Ecol Appl 17:2145–2151.

Miller, S. J. and D. H. Wardrop. 2006. Adapting the Floristic Quality Assessment Index to reflect Anthropogenic Disturbance in Central Pennsylvania Wetlands. Ecological Indicators 6:313–26.

Miller, S.J., D.H. Wardrop, W.M. Mahaney, R. P. Brooks. 2006. A plant-based index of biological integrity (IBI) for headwater wetlands in central Pennsylvania. Ecological Indicators 6:290–312.

Mita, D., E. DeKeyser, D. Kirby, and G. Easson. 2007. Developing a wetland condition prediction model using landscape structure variability. Wetlands 27:1124–1133.

Mitchell, B. R., G.L. Tierney, K.M. Miller, B. Schweiger, D. Faber-Langendoen, J.B. Grace. 2011 (in press). Getting the message across: Using ecological integrity to communicate with resource managers. Pp. xx-xx *in* Guntenspergen. G. et al. Application of threshold concepts in natural resource decision making. Springer, NY.

NatureServe. 2002. Element Occurrence Data Standard. Online at www.natureserve.org/prodServices/eodata.jsp

NatureServe. 2005. NPS PLOTS Database Version 2.0. NatureServe, Arlington, Virginia.

NatureServe. 2008. International Ecological Classification Standard: Terrestrial Ecological Classifications. NatureServe Central Databases. Arlington, VA.

NatureServe. 2009. NatureServe National Systems Map. [N.B. Reference incomplete]

NatureServe Vista version 2.0. 2009. www.natureserve.org/vista. NatureServe, Arlington, VA.

NatureServe. 2011. Central Basin and Range. Rapid Ecoregional Assessments, Final Memorandum I-3-C (March 4, 2011). Department of the Interior, Bureau of Land Management, Rapid Ecoregional Assessments, Denver, Colorado. NatureServe, Washington, DC.

Noon, B. R. 2003. Conceptual issues in monitoring ecological systems. Pp. 27–71 *in* D. E. Busch and J. C. Trexler, editors. Monitoring ecosystems: Interdisciplinary approaches for evaluating ecoregional initiatives. Island Press, Washington, DC. 447 pp.

Oakley, K.L., L P. Thomas, and S G. Fancy. 2003. Guidelines for long-term monitoring protocols. Wildlife Society Bulletin 31(4):1000–1003

Oliver, I., H. Jones, and D.L. Schmoldt. 2007. Expert panel assessment of attributes for natural variability benchmarks for biodiversity. Austral Ecology 32: 453–475.

Parkes, D., G. Newell and D. Cheal. Assessing the quality of native vegetation:The 'habitat hectares' approach ECOLOGICAL MANAGEMENT & RESTORATION VOL 4 SUPPLEMENT: 2003 S29–S38.

Parrish, J.D., D. P. Braun, and R.S. Unnasch. 2003. Are we conserving what we say we are? Measuring ecological integrity within protected areas. BioScience 53: 851–860.

Peet, R. K., T. R. Wentworth, and P. S. White. 1998. The North Carolina Vegetation Survey protocol: a flexible, multipurpose method for recording vegetation composition and structure. Castanea 63:262–274.

Pellant, M., P. Shaver, D.A. Pyke, and J.E. Herrick. 2005. Interpreting indicators of rangeland health, version 4. Technical Reference 1734-6. U.S. Department of the Interior, Bureau of Land Management, National Science and Technology Center, Denver, CO. BLM/WO/ST-00/001+1734/REV05. 122 pp.

Rocchio, J. 2007. Assessing Ecological Condition of Headwater Wetlands in the Southern Rocky Mountain Ecoregion Using a Vegetation Index of Biotic Integrity. Colorado Department of Natural Resources, and U.S. Environmental Protection Agency, Region VIII. Colorado Natural Heritage Program, Colorado State University, Fort Collins, CO. Online: *www.cnhp.colostate.edu/reports.html*

Rowe, J.S. and B.V. Barnes. 1994. Geo-ecosystems and bio-ecosystems. Bulletin of the Ecological Society of America 75 (1): 40–41.

Society for Ecological Restoration (SER) International Science & Policy Working Group. 2004. *The SER International Primer on Ecological Restoration*. www.ser.org & Tucson: Society for Ecological Restoration International. 13 p.

Smith, R. D., A. Amman, C. Bartoldus, and M. M. Brinson. 1995. An approach for assessing wetland functions using hydrogeomorphic classification, reference wetlands, and functional indices. Technical report TR WRP-DE-10, and operational draft. US Army Engineers Waterways Experiment Station, Vicksburg MS.

Stein, B.A., L.S. Kutner, and J.S. Adams (eds.). 2000. Precious Heritage: The status of biodiversity in the United States. The Nature Conservancy and Association for Biodiversity Information [*now NatureServe*]. Oxford University Press.

Stein, E.D., A. E. Fetscher, R. P. Clark, A. Wiskind, J. L. Grenier, M. Sutula, J. N. Collins, and C. Grosso. 2009. Validation of a wetland rapid assessment method: use of EPA's level 1-2-3 framework for method testing and refinement wetlands. Wetlands 29: 648–665.

Stevens, D.L. and S.F. Jensen. 2007. Sample design, execution, and analysis for wetland assessment. Wetlands 27: 515–523.

Stoddard, J.L., D.P. Larsen, C.P. Hawkins, R.K. Johnson, and R.H. Norris. 2006. Setting expectation for the ecological condition of streams: the concept of reference condition. Ecological Applications 16(4): 1267–1276.

Sutula, M.A., E.D. Stein, J.N. Collins, A.E. Fetscher, and R. Clark. 2006. A practical guide for development of a wetland assessment method: the California experience. J. Amer. Water Resources Association 42:157–175.

Swetnam, T.W., C.D. Allen, and J.L. Betancourt. 1999. Applied historical ecology: using the past to manage for the future. Ecological Applications 9:1189–1206.

Tierney, G.L, D. Faber-Langendoen, B.R. Mitchell. G. Shriver, J. Gibbs. 2009. Monitoring and evaluating the ecological integrity of forest ecosystems. Frontiers in Ecology and the Environment 7: 308–316.

Tierney, G. L., B. R. Mitchell, K. M. Miller, J. Comiskey, A. Kozlowski, and D. Faber-Langendoen. 2010. Northeast Temperate Network long-term forest monitoring protocol: 2010 revision., National Park Service, Fort Collins, CO.

Tiner, R.W. 2004. Remotely-sensed indicators for monitoring the general condition of "natural habitat" in watersheds: an application for Delaware's Nanticoke River watershed Ecological Indicators 4 (2004) 227–243.

Unnasch, R.S., D. P. Braun, P. J. Comer, G. E. Eckert. 2009. The Ecological Integrity Assessment Framework: A Framework for Assessing the Ecological Integrity of Biological and Ecological Resources of the National Park System. Report to the National Park Service.

U.S. Department of Agriculture, Natural Resources Conservation Service (USDA NRCS). 1997. National range and pasture handbook. U.S. Department of Agriculture, Washington, DC.

U.S. EPA. 2002. Methods for evaluating wetland condition: wetlands classification. Office of Water, U.S. Environmental Protection Agency, Washington, DC. EPA-822-R-02-017.

U.S. EPA. 2006. Application of Elements of a State Water Monitoring and Assessment Program For Wetlands, Wetlands Division, Office of Wetlands, Oceans and Watersheds, U.S. Environmental Protection Agency, Washington, DC.

USEPA. 2007. Level III Ecoregions of the conterminous United States. *www.epa.gov/wed/pages/ecoregions/level_iii.htm*

USEPA (U.S. Environmental Protection Agency). 2011. National Wetland Condition Assessment: Field Operations Manual. EPA-843-R-10-001. U.S. Environmental Protection Agency, Washington, DC.

U.S. Fish and Wildlife Service (USFWS). 2010. NWIPlus: Geospatial Database for Watershed-level Functional Assessment. April 2010 Factsheet. *http://www.fws.gov/northeast/wetlands/factsheets/NWIPlus_FactSheet.pdf*

Voss, E. G. 1972. Michigan Flora. Part I. Gymnosperms and Monocots. Cranbrook Institute of Science Bulletin 55 and the University of Michigan Herbarium. xv + 488 pp.

Voss, E. G. 1985. Michigan Flora. Part II. Dicots (Saururaceae-Cornaceae). Cranbrook Institute of Science Bulletin 59 and the University of Michigan Herbarium. xix + 724 pp.

Voss, E. G. 1996. Michigan Flora. Part III. Dicots (Pyrolaceae-Compositae). Cranbrook Institute of Science Bulletin 61 and the University of Michigan Herbarium. xix + 622 pp.

Williams, J.W., S.T. Jackson, and J. Kutzbach. 2007. Projected distributions of novel and disappearing climates by 2100 AD. PNAS 104: 5738–5742

Yorks, T. E., and S. Dabydeen. 1998. Modification of the Whittaker sampling technique to assess plant diversity in forested natural areas. Natural Areas Journal 18:185–189.

Young, T.F. and S. Sanzone (editors). 2002. A framework for assessing and reporting on ecological condition. Prepared by the Ecological Reporting Panel, Ecological Processes and Effects Committee. EPA Science Advisory Board. Washington, DC. 142 pp.